

# Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification

Yusi Fu<sup>a</sup>, Chunmei Li<sup>a</sup>, Sijia Lu<sup>b,1</sup>, Wenxiong Zhou<sup>a</sup>, Fuchou Tang<sup>a,c</sup>, X. Sunney Xie<sup>a,b,2</sup>, and Yanyi Huang<sup>a,c,d,2</sup>

<sup>a</sup>Biodynamic Optical Imaging Center, School of Life Sciences, Peking University, Beijing 100871, China; <sup>b</sup>Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138; <sup>c</sup>Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China; and <sup>d</sup>College of Engineering, Peking University, Beijing 100871, China

Contributed by X. Sunney Xie, July 28, 2015 (sent for review June 7, 2015; reviewed by Luke P. Lee)

**Whole-genome amplification (WGA) for next-generation sequencing has seen wide applications in biology and medicine when characterization of the genome of a single cell is required. High uniformity and fidelity of WGA is needed to accurately determine genomic variations, such as copy number variations (CNVs) and single-nucleotide variations (SNVs). Prevailing WGA methods have been limited by fluctuation of the amplification yield along the genome, as well as false-positive and -negative errors for SNV identification. Here, we report emulsion WGA (eWGA) to overcome these problems. We divide single-cell genomic DNA into a large number (10<sup>5</sup>) of picoliter aqueous droplets in oil. Containing only a few DNA fragments, each droplet is led to reach saturation of DNA amplification before demulsification such that the differences in amplification gain among the fragments are minimized. We demonstrate the proof-of-principle of eWGA with multiple displacement amplification (MDA), a popular WGA method. This easy-to-operate approach enables simultaneous detection of CNVs and SNVs in an individual human cell, exhibiting significantly improved amplification evenness and accuracy.**

single cell | whole-genome amplification | sequencing | microfluidics | emulsion

**S**ingle-cell sequencing, characterization the genome of individual cells, is highly needed for studying scarce and/or precious cells, which are inaccessible for conventional bulk genome characterization, and for probing genomic variations of a heterogeneous population of cells (1–3). Recently single-cell genomics has unveiled unprecedented details of various biological processes, such as tumor evolution (4–6), embryonic development (7), and neural somatic mosaicism (8). Single-cell whole-genome amplification (WGA) is required to generate enough replicates of genomic DNAs for library preparation in conjunction with current sequencing protocols. Single-cell WGA has been increasingly used in cutting-edge clinical diagnostic applications such as molecular subtyping of single tumor cells (4, 9) and preimplantation genetic screening of in vitro fertilized embryos (10).

An ideal single-cell WGA method should have high uniformity and accuracy across the whole genome. The WGA uniformity is critical for copy number variation (CNV) detection, whereas the WGA accuracy is essential for avoiding single-nucleotide variation (SNV) detection errors, either false positives or false negatives. The false positives arise from misincorporation of wrong bases in the first few cycles of WGA. In a diploid human cell, the false negatives primarily arise from the allelic dropout (ADO), i.e., heterozygous mutations are mistaken as homozygous ones because of the lack of amplification in one of the two alleles (11).

Existing WGA chemistry includes degenerate oligonucleotide-primed PCR (DOP-PCR) (12), multiple displacement amplification (MDA) (13–17), and multiple annealing and looping-based amplification cycles (MALBAC) (4, 18, 19), which have successively achieved genome analysis at the single-cell level. DOP-PCR is based on PCR amplification of the fragments flanked by universal priming sites, and provides high accuracy for detecting

CNVs in single cells but has low coverage and high false-positive and false-negative rates for calling SNVs (5). MDA has a much improved coverage but tends to have lower precision/sensitivity in CNV determination due to its variation of the amplification gain along the genome, not reproducible from cell to cell (20). By virtue of quasilinear amplification, MALBAC suppresses the random bias of amplification and exhibits reduced ADO rates, yielding low false negatives for SNV detection (2, 11, 18, 19). Notwithstanding its drawbacks, MDA still offers comparable or higher genome coverage than MALBAC, at least for single diploid cells, possibly taking advantage of the randomness (2). In fact, even higher coverage has been obtained for cells with aneuploidy, such as dividing cells (21), and cancer cells (22). MDA's main advantage is its lower false-positive rate for SNV detection on account of the use of Phi-29, a highly processive polymerase with high fidelity.

Microfluidic devices have been carried out for single-cell WGA (16, 20, 23, 24), allowing avoidance of contaminations and high-throughput analyses of multiple single cells in parallel. The small total reaction volumes (microliters to nanoliters- or picoliters) of the microfluidic devices not only facilitate the efficiency of reactions but also allow significant cost reduction for enzymes and reagents used. It was reported that the nanoliter volume of a microfluidic device improved uniformity of the amplification compared with microliter devices in the WGA of single bacterial cells (20).

Here, we report a method, emulsion whole-genome amplification (eWGA), to use the small volume of aqueous droplets in oil to better the WGA chemistry for uniform amplification of a single

## Significance

**Uniform and accurate single-cell whole-genome amplification is important when starting material is limited and precious. We develop an emulsion-based amplification method that can suppress the amplification bias to detect high-resolution copy number variations of a single cell, and to simultaneously detect the single-nucleotide variations with high accuracy. This approach is compatible with various amplification protocols including the widely used multiple displacement amplification, which has been demonstrated in this paper.**

Author contributions: Y.H. designed research; C.L. prepared sequencing libraries; Y.F., C.L., and W.Z. performed research; Y.F., S.L., F.T., X.S.X., and Y.H. analyzed data; and Y.F., S.L., F.T., X.S.X., and Y.H. wrote the paper.

Reviewers included: L.P.L., University of California, Berkeley.

Conflict of interest statement: S.L. and X.S.X. are cofounders and shareholders of Yikon Genomics.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the NCBI Sequence Read Archive database (accession no. [SRP052908](https://www.ncbi.nlm.nih.gov/sra/SRP052908)).

<sup>1</sup>Present address: Yikon Genomics, Co., Ltd., Taizhou, Jiangsu 225300, China.

<sup>2</sup>To whom correspondence may be addressed. Email: [xie@chemistry.harvard.edu](mailto:xie@chemistry.harvard.edu) or [yanyi@pku.edu.cn](mailto:yanyi@pku.edu.cn).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1513988112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1513988112/-DCSupplemental).

cell's genome. By distributing single-cell genomic DNA fragments into a large number ( $10^5$ ) of picoliter droplets, a few DNA fragments in each droplet is allowed to reach saturation of DNA amplification. After merging the droplets by demulsification, the differences in amplification gain among the DNA fragments are significantly minimized.

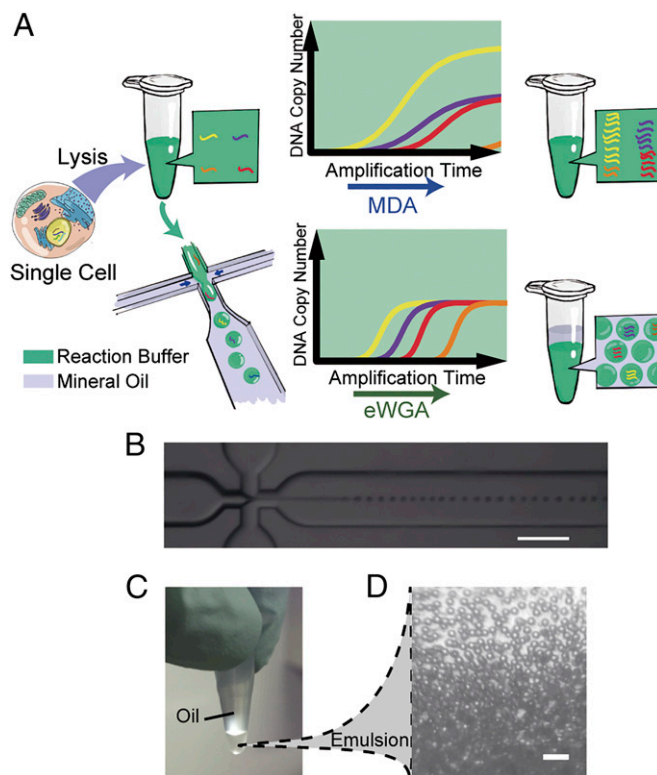
Although this approach can be used for any chemistry of WGA, we take MDA as an example to greatly reduce the random bias of amplification by separating the reactions into a large amount of emulsion droplets. We carried out detailed comparison with MDA, MALBAC, and DOP-PCR performed in tube using single cells from normal diploid human cells and a monoclonal human cancer cell line with inherited CNVs. Our results indicate that eWGA not only offers higher coverage but also enables simultaneous detection of SNVs and CNVs with higher accuracy and finer resolution, outperforming the prevailing single-cell amplification methods in many aspects.

## Results and Discussion

**eMDA Sequencing Library Preparation.** MDA, an easy-operating and widely used single-cell WGA protocol, is used for the proof-of-concept of eWGA. We lysed individual cells to release the genomic DNA (gDNA) fragments and dehybridized them to single strands by heating. After adding the MDA reaction buffer, the solution ( $10 \mu\text{L}$ ) was distributed into  $\sim 7 \times 10^5$  droplets,  $14 \text{ pL}$  each, using a microfluidic chip (Fig. 1 A and B, and *SI Appendix, Fig. S1*). This process is carried out at  $4^\circ\text{C}$  to keep the amplification from starting. Under this lysing condition, the estimated mean size of DNA fragments is  $\sim 10 \text{ kb}$  (18). Thus, for a single diploid cell, each droplet contains one fragment on average. We have tested different dilutions of DNA and observed the decline of mapping rate with further dilution, especially when the average fragment is far less than one per droplet. This is because a large number of empty droplets increases the ratio of nonspecific product of amplification. On the other hand, more DNA fragments in one droplet ( $>10$  per droplet) impairs the evenness of WGA. The aneuploidy of the cell will affect the actual number of fragments per droplet. We found eMDA performance is stable when each droplet has one to two fragments. We collected all of the droplets in a microcentrifuge tube (Fig. 1 C and D). In contrast to the conventional single-tube MDA reaction, which exhibits more serious amplification bias with longer reaction time, in the emulsion MDA (eMDA) reaction each droplet produces similar amount of amplification products due to the eventual saturation of the polymerization reaction in each droplet (*SI Appendix, Fig. S2*). After heat inactivation of the enzyme and demulsification, the amplification uniformity is accomplished in the aqueous solution, and the amplification products are used to construct sequencing libraries.

### eMDA Amplify Normal Diploid Single Cells Evenly and Completely.

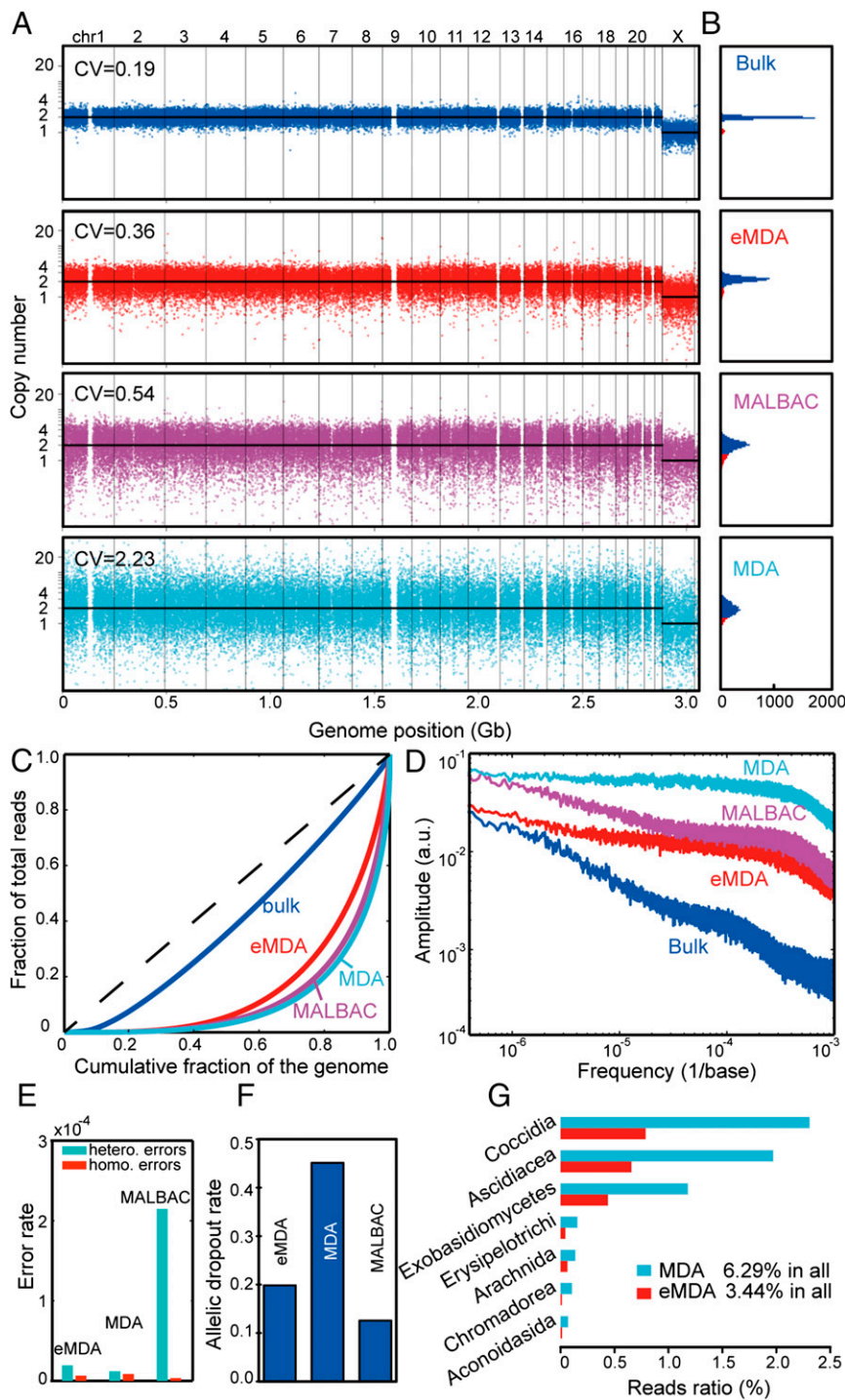
We chose human umbilical vein endothelial cell (HUVEC), a normal human diploid cell line, to validate the amplification evenness of eMDA using bulk ( $200 \text{ ng}$ ) genomic DNA from HUVECs as a reference. We carried out 10 single-cell eMDA experiments and compared the sequencing results with those of single-cell MALBAC or conventional MDA reactions. We divided the human genome into bins with mean size of  $52.4 \text{ kb}$  using dynamic binning method (5) and applied shallowly sequenced data,  $3\text{M}$  uniquely mapped reads for each single cell, to calculate the copy number in each bin (Fig. 2A and *SI Appendix, Figs. S3 and S4*). eMDA showed the most uniform amplification across the whole genome, with coefficient of variation (CV) of 0.36, which is significantly lower than the conventional MDA ( $\text{CV} = 2.23$ ) (Fig. 2A, Table 1, and *SI Appendix, Fig. S5*). From the reads covering autosomes and sex chromosomes of these single HUVECs (Fig. 2B), we found eMDA providing the smallest deviation from a priori expectation using bulk DNA as a reference.



**Fig. 1.** The experimental process of eWGA-seq and emulsion generation. (A) A single cell is lysed and then mixed with MDA reaction buffer in a tube. The solution was either directly used for conventional MDA, generating unevenly amplified DNA fragments, or used for emulsion generation in a microfluidics cross-junction device, resulting in uniformly distributed aqueous reaction droplets and evenly amplified DNA fragments. (B) The microfluidics cross-junction. Reaction buffer and mineral oil are driven by compressed air with proper pressure to achieve uniform water-in-oil emulsion. The cross-section of the channel is  $105 \times 100 \mu\text{m}$ . The speed of emulsion generation is  $\sim 35,000$  per min. (Scale bar:  $300 \mu\text{m}$ .) (C) All droplets are collected into a  $200\text{-}\mu\text{L}$  microcentrifuge tube and incubated at  $30^\circ\text{C}$  to perform eWGA. (D) The emulsion is stable during the reaction. (Scale bar:  $100 \mu\text{m}$ .)

We sequenced a few single HUVECs to a greater depth ( $>14\times$ ) using eMDA, MALBAC, and conventional MDA, and plotted the Lorenz curves of coverage to further validate the evenness of eMDA (Fig. 2C). As perfectly uniform coverage would result in a diagonal line, eMDA shows the best uniformity across the whole genome, compared with MALBAC and conventional MDA-amplified single cells, and is closest to the unamplified bulk sample. In contrast to the previously reported nanoliter MDA reaction in which the amplification gain is reduced (20, 23), our eMDA yields a similar gain as the conventional MDA to ensure a high coverage breadth of the genome. We showed that emulsion would not result in losing fragments of DNA as eMDA exhibits slightly higher coverage breadth (72.3% at  $10\times$ , for a human diploid cell) than MALBAC (67.5%) or conventional MDA (68.5%) at the same sequencing depth (*SI Appendix, Figs. S6 and S7*).

We also plotted the power spectra of read density as a function of the spatial frequency (Fig. 2D) based on the sequencing result using different protocols. The analysis confirmed that, for single-cell sequencing, eMDA provides the best uniformity among the three methods by offering smaller copy number fluctuation at all frequencies due to the effectively suppressed amplification bias through compartmentation. Because the intrinsic amplification randomness still exists within each droplet, the uniformity improvement is more significant in the lower frequency (large bin size) region than in the higher frequency domain.



**Fig. 2.** The comparison of WGA methods for sequencing single HUVECs. (A) The copy number across the whole genome with a mean bin size of 52.4 kb; black line shows the expected value. (B) The density histogram of copy number distribution (bin size, 502 kb). (C) The Lorenz curves of coverage uniformity for single cells amplified by eMDA, MALBAC, conventional MDA, and unamplified genomic DNA. (D) The power spectrum of read density as a function of spatial frequency. (E) Copying-error rate of single-cell WGA methods. (F) ADO rate of single-cell WGA methods. (G) The ratio of the sequencing read originated from major pollutes in single-cell eMDA and conventional MDA experiments.

To estimate the accuracy of CNV identification of these methods, we carried out a simulation by calling the artificial CNVs with both copy number gain (2 to 3) and loss (2 to 1), in silico generated within diploid autosomes (SI Appendix, Fig. S8). The accuracy is the ratio of simulated CNVs that could be detected at the 52.4-kb resolution. eMDA shows much higher accuracy to identify the CNVs at the range from 300 kb to 2 Mb. We also performed an

intersample correction (24) for MALBAC to eliminate the sequence-dependent bias (SI Appendix, Fig. S9), whereas for eMDA such normalization is unnecessary. This feature is very important in various medical applications such as in vitro fertilization preimplantation screening because a standard normalization sample and the expertise of performing complicated cross-sample normalization are often not available. eMDA was



**Table 1. Summary of the comparison between different methods for single human cell amplification**

Parameter	Amplification method			
	eMDA	MDA	MALBAC*	DOP-PCR
<b>HT-29 single cells</b>				
CNV resolution, kb <sup>†</sup>	619	9,669	5,847	538
CNV accuracy, % <sup>‡</sup>	66.5	50.9	41.3	78.8
False-positive rate, % <sup>§,¶</sup>	0.01	0.02	0.04	0.30
Coverage breadth, % <sup>§</sup>	90.3	74.4	78.8	43.7
SNV false-negative rate, % <sup>§,#</sup>	11.7	39.3	27.2	70.3
ADO rate, % <sup>§,  ,**</sup>	39.3	52.9	23.8	88.5
<b>HUVEC single cells</b>				
CV, 52.4k bin	0.45	2.23	0.55	—
CNV resolution, kb, gain/loss <sup>†,††</sup>	1,150/350	>2,000/1,200	>2,000/700	—
False-positive rate, % <sup>¶,†††</sup>	0.08	0.09	0.12	—
Coverage breadth, % <sup>††</sup>	72.3	68.5	67.5	—
SNV false-negative rate, % <sup>#</sup>	29.5	30.6	34.6	—
ADO rate, % <sup>  ,**</sup>	19.8	45.1	12.6	—
Error rate, % <sup>§§</sup>	0.0019	0.0012	0.021	—

\*The MALBAC data have been corrected through intersample normalization.  
<sup>†</sup>The sensitivity cutoff is set to be 90%.  
<sup>‡</sup>CNVs that are larger than 500 kb.  
<sup>§</sup>Calculation based on exome-enriched sequencing data.  
<sup>¶</sup>False-positive rate (FPR) is higher for the whole-genome sequencing data compared with exome sequencing data due to the relative lower coverage depth of some loci in whole-genome sequencing data.  
<sup>#</sup>SNV false-negative rate is lower for HT-29 single cells than for HUVECs because of the hyperploidy of HT-29 cancer cells.  
<sup>||</sup>Calculation was based on the diploid region with coverage depth larger than 30x.  
<sup>\*\*</sup>Allelic dropout rate (ADO) calculated from exome sequencing data is larger than that from whole-genome sequencing data due to the material loss during capture.  
<sup>††</sup>Calculation was based on simulation of copy number gains and losses ranging from 250 kb to 2 Mb.  
<sup>†††</sup>Calculation was based on 30-Gb sequencing data of each method.  
<sup>§§</sup>Calculation was based on reads mapped on chromosome X.

superior to both MALBAC and conventional MDA by offering finer smallest detectable CNV events (350 kb for copy number loss and 1.2 Mb for copy number gain, at 90% sensitivity in the diploid genomic region).

**eMDA Amplifies Normal Diploid Single Cells with Higher Accuracy.**

From the deeply sequenced single-cell data, we detected more homozygous and heterozygous SNVs by eMDA than by MALBAC or conventional MDA (SI Appendix, Table S1), in accordance with the higher coverage breadth. As the HUVEC cells we used were from a male, we then deduced the error rates of these methods by calculating the ratio between high-confidence heterozygous SNVs and homozygous SNVs on the X chromosome from each dataset. The error rate of eMDA ( $1.9 \times 10^{-5}$ ) was comparable with that of conventional MDA ( $1.2 \times 10^{-5}$ ), but one order of magnitude less than that of MALBAC ( $2.1 \times 10^{-4}$ ) (Fig. 2E). These values, which matched well with previous reports (20), faithfully reflected the difference between the high fidelity Phi-29 polymerase used in eMDA and MDA, and the error-prone enzyme used in MALBAC which lacks proofreading capability.

We then examined the ADO rate of these methods by identifying the loss-of-heterozygosity events in the high-confidence heterogeneous SNVs (>20x coverage depth and >20% for each allele) found in autosomes from the bulk. For a normal diploid HUVEC, the ADO rate of eMDA is 19.8% (Fig. 2F). This performance is close to MALBAC, with which the ADO rate is ~12%, making eMDA a great choice for those single-cell applications that could not be implemented by conventional MDA due to its notoriously high ADO rate (45.1%).

MDA is prone to environmental contamination including the trace amount of DNA pollution in reagents. The contaminant DNA could be reduced by applying small reaction volumes (20, 25). With eMDA, the reaction buffer is distributed to a large number of separated droplets, and the contaminant DNA will only exist in a small portion of droplets and not be overamplified. In addition, because the single human cells are carefully picked through micromanipulation under a microscope, and washed multiple times before lysing, the contamination from other mammalian cells is minimized. Metagenomic analysis (Fig. 2G) verified that eMDA produced much cleaner (3.4% nonhuman reads) data than MDA did (6.3% nonhuman reads) for single HUVEC sequencing.

**High-Resolution Inherited CNV Detection in Single Cancer Cells.**

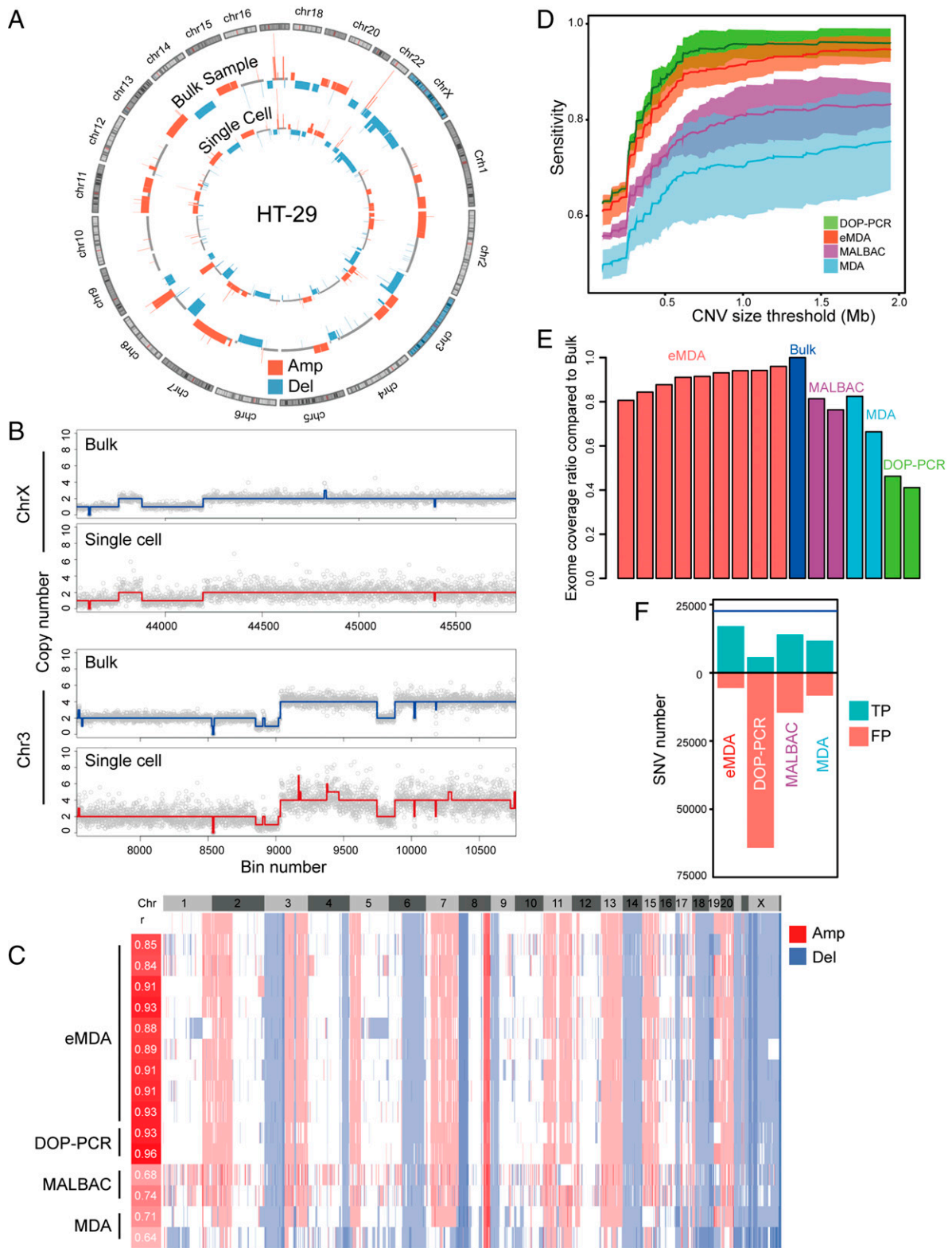
We next applied eMDA to sequence nine single HT-29 cancer cells expanded from a single clone. HT-29 is a colon adenocarcinoma cell line with multiple chromosomal aberrations, making its nuclear DNA close to triploid (26). We validated the aneuploidy through flow cytometry (SI Appendix, Fig. S10) and observed that the coverage depth pattern (27) of each single cell is similar to that of bulk (200 ng) gDNA (Fig. 3A). We called the CNVs from eMDA-amplified single cells at different resolutions, and found that the CNV pattern of each single cell is almost identical to that of the monoclonal expanded bulk sample, with correlation  $r = 0.90 \pm 0.03$ ,  $0.95 \pm 0.02$ , and  $0.96 \pm 0.02$  at 52.4-kb, 502-kb, and 5-Mb resolution, respectively. At the 52.4-kb resolution, we were able to identify CNVs with smallest size of ~250 kb, which was the 5-bin cutoff we applied to the analysis (Fig. 3B). We also profiled CNV patterns of single cells amplified from MALBAC, DOP-PCR, or conventional MDA at 52.4-kb resolution (SI Appendix, Fig. S11) and found that, compared with MALBAC and conventional MDA, the improved amplification uniformity of eMDA allowed us to obtain more reliable genomewide CNV pattern (Fig. 3C) as well as the higher specificity and higher sensitivity of CNV identification in single cells, with performance close to DOP-PCR (Fig. 3D and SI Appendix, Fig. S12A).

**Exome Coverage Breadth and SNV Detection in Single Cancer Cells.**

We then investigated the accuracy of SNV identification from single HT-29 cells using eMDA. We performed exome enrichment and sequencing for all samples and used bulk HT-29 exome as a reference. eMDA shows highest coverage ( $\geq 1 \times$  depth,  $90 \pm 5\%$ ), followed by MALBAC ( $79 \pm 4\%$ ), conventional MDA ( $74 \pm 11\%$ ), and DOP-PCR ( $44 \pm 4\%$ ) (Fig. 3E). eMDA also exhibits high accuracy to identify homozygous SNVs of single cells, with highest true-positive ratio and lowest false-positive rate among all methods we tested (Fig. 3F). As expected, eMDA also noticeably reduce the ADO to 24% from 43% of conventional MDA for these nondiploid single cells (SI Appendix, Fig. S12B).

**Conclusion**

Our method, eWGA, applies emulsion to divide the DNA fragments from a single cell to a large number of aqueous droplets in oil and drives the amplification to saturation in each droplet. Using MDA protocol as a demonstration, this approach can dramatically reduce the amplification bias while retaining the high accuracy of replication. Unlike other microfluidics-based WGA methods (20, 23), which improved the uniformity by reducing



**Fig. 3.** The comparison of WGA methods for sequencing single HT-29 cells. (A) The circos plot (27) showing the copy number profiles from unamplified genomic DNA and from a single cell amplified by eMDA. (B) The zoomed-in copy number distribution of chr3 and chrX with a binning size of 52.4 kb. The smallest CNV detected is 5 bins. (C) Heat map showing copy number gains and losses of single cells with different amplification methods, with unamplified genomic DNA as reference. The correlation efficiencies between single-cell WGA methods and bulk reference are also listed. (D) The CNV detection sensitivity under different bin size threshold of single-cell WGA methods. The filled area represents the SD of each method. (E) The coverage ratio of exome captured single-cell WGA samples using unamplified sample as reference. (F) The homozygous SNVs detected in single cells using different WGA methods. The blue line shows the number of homozygous SNVs identified in the unamplified sample. The blue bars show the SNVs that matched bulk reference, whereas the red bars show the discordant SNVs.

the gain compared with conventional MDA, eMDA has the gain of  $\sim 2 \times 10^6$ , which is comparable to the conventional MDA in tube with single human cells as starting material. With the high coverage breadth across the whole genome, eMDA also enables us to detect more SNVs than existing methods and the pollution rate is alleviated with the use of emulsion. eMDA is compatible with targeted enrichment methods such as exome capture, which is useful when only certain regions are of interest in genetic analyses. By using eMDA, the first method (to our knowledge) that enables simultaneous identification of both small CNVs and high-confidence SNVs from a single human cell, we are able to detect CNVs at 250-kb size with 50-kb resolution, and SNVs with error rate  $< 2 \times 10^{-5}$ . We envision that such emulsion approach will also improve the amplification performance of other WGA methods, for example MALBAC, for single-cell genomic studies.

## Materials and Methods

**Device Fabrication.** Microfluidic emulsion-generating chips were made of polydimethylsiloxane (PDMS). The mold used to cast the chips was made by etching photoresist on a silicon wafer using photolithography. In brief, SU-8 2025 (MicroChem) was spin coated onto the wafer at 1700 rpm for 60 s on a spin coater (KW-4A, SETCAS Electronics Co., Ltd), resulting in a thickness of 50  $\mu\text{m}$  of photoresist. Then the wafer was baked at 95  $^{\circ}\text{C}$  for 5 min. The wafer was exposed to UV light for 30 s through a mask defining the channel geometry and then the wafer was baked again at 95  $^{\circ}\text{C}$  for 10 min. The unexposed photoresist was removed with solvent and the wafer was hard-baked at 150  $^{\circ}\text{C}$  for 3 h. The mold was treated with trimethyl chlorosilane vapor for 10 min before use. Then 30 g degassed and well-mixed 5:1 (base:curing agent) PDMS (Sylgard 184, Dow Corning) was poured on the wafer, and baked together at 80  $^{\circ}\text{C}$  for 15 min before peeled off. Then we punched the holes for the inlets of reagent/oil and the outlet for connecting a micro-tubing that transferred the emulsion droplets to a 200  $\mu\text{L}$  micro-centrifuge tube. Then the patterned PDMS

slab was bonded with a piece of cover glass precoated with 20:1 (base:curing agent) PDMS through baking at 80  $^{\circ}\text{C}$  for 3 h. The resulting chip is shown as [SI Appendix, Fig. S1](#).

**Conventional Single-Cell MDA Reaction.** The gDNA was fragmented by heating (4 min at 98  $^{\circ}\text{C}$ , and 2 min at 95  $^{\circ}\text{C}$ ) in 4  $\mu\text{L}$  lysis buffer [30 mM Tris-HCl (pH = 8.0), 10 mM NaCl, 1 mg/mL proteinase (Qiagen), 5 mM EDTA and 0.5% Triton X-100]. Then 6  $\mu\text{L}$  MDA reaction buffer was added to reach 10  $\mu\text{L}$  total volume with a final concentration of 1x Phi-29 buffer (NEB), 50  $\mu\text{M}$  N6 primer with two phosphorothioate bonds at the 3'-side (Invitrogen), 1 mM dNTP (NEB), 0.2 mg/mL BSA (NEB). We heated the tube at 95  $^{\circ}\text{C}$  for 5 min, and then immediately put it on ice for at least 20 min to anneal the random hexmers to fragmented gDNA. We then added 8 units of Phi-29 polymerase (NEB) and briefly centrifuged. Then MDA reactions were carried out at 30  $^{\circ}\text{C}$ . Reactions were terminated at 65  $^{\circ}\text{C}$  for 10 min after 10 h amplification.

**Single-Cell eMDA Reaction.** The reaction buffer preparation is identical to MDA reactions. However, to prevent the reaction from initiating prior to droplet generation, the Phi-29 polymerase was added to the reaction mix immediately before emulsion generation. The reaction buffer was kept at 4  $^{\circ}\text{C}$  to prevent the amplification from starting before being dispersed into droplets. The emulsion droplets were collected into a tube and then incubated at 30  $^{\circ}\text{C}$  for 8–10 h before termination at 65  $^{\circ}\text{C}$  for 10 min.

A detailed description of remaining material and methods can be found in [SI Appendix](#).

**ACKNOWLEDGMENTS.** We thank Dr. Ruiqiang Li, Dr. Zhilong Yu, Xiannian Zhang, Wei Fan, Ang Li, Zitian Chen, Tao Chen, and Zhe Su for their help on experiments and data analysis, Dr. Yun Zhang and the High-Throughput Sequencing Center of Peking University for assistance with Illumina sequencing, and the National Center for Protein Sciences Beijing (Peking University) for assistance with FACS analysis. This work was supported by National Natural Science Foundation of China Grants 21327808, 91313302, and 21222501 (to Y.H.).

- Kalisky T, Quake SR (2011) Single-cell genomics. *Nat Methods* 8(4):311–314.
- Huang L, Ma F, Chapman A, Lu S, Xie XS (2015) Single-cell whole-genome amplification and sequencing: Methodology and applications. *Annu Rev Genom Human Genet*, 10.1146/annurev-genom-090413-025352.
- Wang Y, Navin NE (2015) Advances and applications of single-cell sequencing technologies. *Mol Cell* 58(4):598–609.
- Ni X, et al. (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci USA* 110(52):21083–21088.
- Navin N, et al. (2011) Tumour evolution inferred by single-cell sequencing. *Nature* 472(7341):90–94.
- Vogelstein B, et al. (2013) Cancer genome landscapes. *Science* 339(6127):1546–1558.
- Yan L, et al. (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* 20(9):1131–1139.
- McConnell MJ, et al. (2013) Mosaic copy number variation in human neurons. *Science* 342(6158):632–637.
- Klein CA, et al. (2002) Combined transcriptome and genome analysis of single micrometastatic cells. *Nat Biotechnol* 20(4):387–392.
- Hou Y, et al. (2013) Genome analyses of single human oocytes. *Cell* 155(7):1492–1506.
- Navin NE (2014) Cancer genomics: One cell at a time. *Genome Biol* 15(8):452.
- Telenius H, et al. (1992) Degenerate oligonucleotide-primed PCR: General amplification of target DNA by a single degenerate primer. *Genomics* 13(3):718–725.
- Dean FB, Nelson JR, Giesler TL, Lasken RS (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res* 11(6):1095–1099.
- Zhang K, et al. (2006) Sequencing genomes from single cells by polymerase cloning. *Nat Biotechnol* 24(6):680–686.
- Chitsaz H, et al. (2011) Efficient de novo assembly of single-cell bacterial genomes from short-read data sets. *Nat Biotechnol* 29(10):915–921.
- Wang J, Fan HC, Behr B, Quake SR (2012) Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell* 150(2):402–412.
- Peters BA, et al. (2012) Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells. *Nature* 487(7406):190–195.
- Zong C, Lu S, Chapman AR, Xie XS (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 338(6114):1622–1626.
- Lu S, et al. (2012) Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science* 338(6114):1627–1630.
- de Bourcy CFA, et al. (2014) A quantitative comparison of single-cell whole genome amplification methods. *PLoS One* 9(8):e105585.
- Wang Y, et al. (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 512(7513):155–160.
- Xu X, et al. (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* 148(5):886–895.
- Gole J, et al. (2013) Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat Biotechnol* 31(12):1126–1132.
- Yu Z, Lu S, Huang Y (2014) Microfluidic whole genome amplification device for single cell sequencing. *Anal Chem* 86(19):9386–9390.
- Blainey PC, Quake SR (2011) Digital MDA for enumeration of total nucleic acid contamination. *Nucleic Acids Res* 39(4):e19.
- Kawai K, et al. (2002) Comprehensive karyotyping of the HT-29 colon adenocarcinoma cell line. *Genes Chromosomes Cancer* 34(1):1–8.
- Krzywinski M, et al. (2009) Circos: An information aesthetic for comparative genomics. *Genome Res* 19(9):1639–1645.