analytical chemistry

Letter

pubs.acs.org/ac

Microfluidic Whole Genome Amplification Device for Single Cell Sequencing

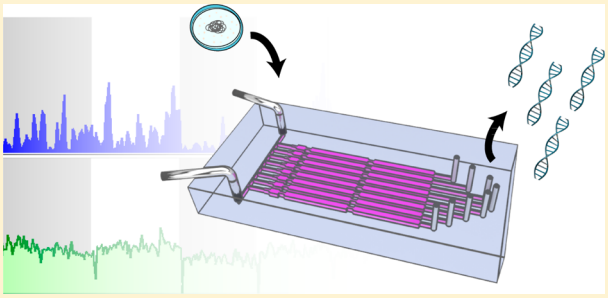
Zhilong Yu,[†] Sijia Lu,[‡] and Yanyi Huang*,[†]

[†]Biodynamic Optical Imaging Center (BIOPIC) and College of Engineering, Peking University, Beijing 100871, China

[‡]Yikon Genomics Co. Ltd., 1 China Medical Ave., Taizhou, Jiangsu 225300, China

Supporting Information

ABSTRACT: We developed a microfluidic device to perform multiplex single-cell whole-genome amplification (WGA) using multiple annealing and looping-based amplification cycles (MALBAC). This device, made of polydimethylsiloxane (PDMS), allows us to monitor the whole process of cell loading and single-cell WGA for sequencing. We show that the genome coverage of MALBAC amplifications is reproducible between chambers on a single chip and between different chips, which enables data normalization using standard samples to accurately identify copy number variations (CNVs). This device provides an easy-to-operate approach to perform single cell sequencing library preparation with minimum hands-on time. It reduces the requirement of manual expertise as well as the risk of contamination, which is essential in future applications especially the medical diagnosis.



High throughput sequencing is becoming one of the most powerful tools to collect complex data for cutting-edge biology and medical research. Single cell sequencing, by investigating the genomic heterogeneity among cells, reveals the hidden information that is often masked by bulk measurement.[1] Single cell genome sequencing enables more delicate analysis on challenging topics including haplotype analysis,[2] cancer genomics,[3] and genomic variability.[4] Research with limited numbers of cells as starting material, such as embryo samples,[5−7] circulating tumor cells,[8,9] and primary neurons,[10] also requires single cell genome sequencing and analysis. While the high-throughput sequencing is getting more and more affordable, the sample preparation, especially the sequencing library construction from single cells, becomes a challenge.

Whole genome amplification (WGA) is required for single cell genome sequencing because a single mammalian cell contains only a few picograms of genomic DNA, which is far from enough for preparing the sequencing library. Several WGA methods have been developed to produce nano- to microgram DNA fragments from a single cell. Among these methods, polymerase chain reaction (PCR)-based WGA (for example, DOP-PCR)[11,12] and multiple displacement amplification (MDA)[13,14] have shown promising results in a variety of applications. However, these methods suffer from significant amplification bias and uneven coverage across the whole genome.[15,16] A recently developed method, multiple annealing and looping-based amplification cycles (MALBAC),[17] has shown improved amplification uniformity across the whole genome and also provides a higher sequencing coverage.

One of the major obstacles in single cell genomic studies is that the amplification process is prone to contamination from exogenous DNA fragments and pre-existed genetic materials. Microfluidic devices that typically deal with nanoliter instead of microliter reactions intrinsically reduce the likelihood of contamination. Such a feature is previously demonstrated by quantitative PCR,[18,19] transcriptome sequencing,[20] and MDA.[21]

Here, we present a novel microfluidic device to perform multiplex single-cell WGA using MALBAC. Up to eight single-cell MALBAC reactions can be performed in parallel using a single device. The whole WGA process, including cell lysis and a two-step MALBAC process containing preamplification and PCR amplification, can be completed within 4 h with minimal hands-on time. The results are reproducible between chambers on a single chip and between different chips. The samples share a similar pattern on sequencing amplification bias, which enable us to perform a robust intersample normalization.

All devices were fabricated using multilayer soft lithography.[22] Each microfluidic chip has eight amplification units, and each unit contains three cascading chambers (Figure 1a): a cell lysis chamber (75 nL), a MALBAC preamplification chamber (500 nL), and a MALBAC PCR chamber (500 nL). The height of the chambers is 50 μm. The width of the common channel is 250 μm, while the width of the channels between the chambers is 420 μm. Additional channels (not shown in the figure) are added around the control channels to decrease sample evaporation during thermocycling. We chose

Received: August 27, 2014
Accepted: September 18, 2014
Published: September 18, 2014

 © 2014 American Chemical Society dx.doi.org/10.1021/ac5032176 | *Anal. Chem.* 2014, 86, 9386−9390

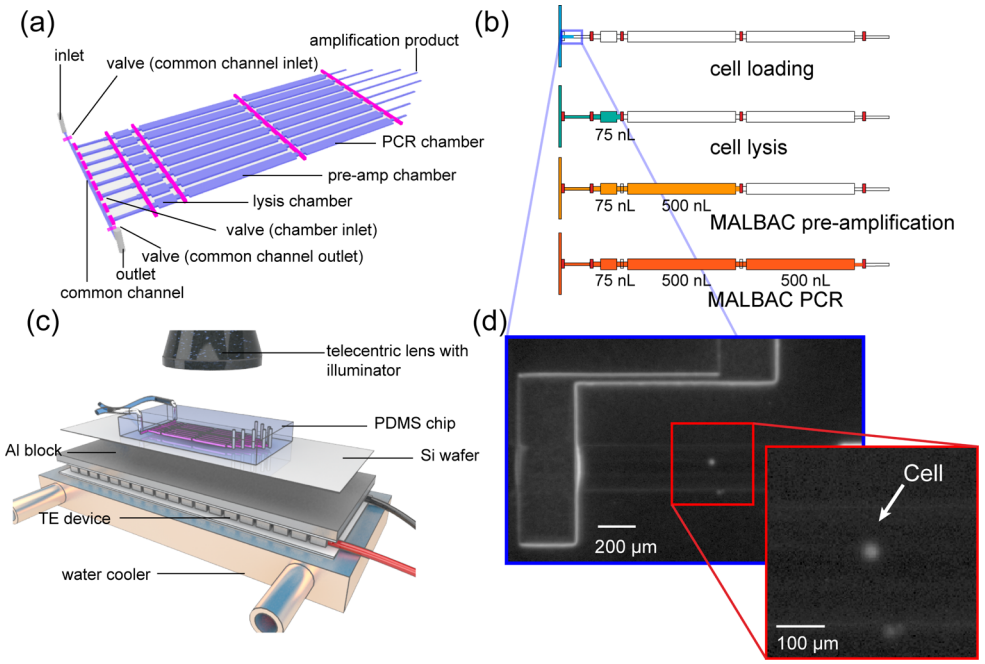


Figure 1. An integrated microfluidic device designed for single-cell MALBAC reactions. (a) Schematics of the device showing the fluidic channels (purple) and the control channels (magenta). (b) The operation of MALBAC reactions on a chip. Cells are suspended in PBS, and single cells are loaded manually by controlling the corresponding valves. Then, the cells are lysed and a two-step MALBAC reaction is performed. (c) Thermocycler and imaging system for the MALBAC reaction device. The PDMS chip is bonded to a Si wafer and placed on a Peltier device, which is attached to a water cooler. (d) Scattering of a single cell on-chip.

mouse embryonic stem cells (mESCs) as our model system through the experiments.

The chip and tubes were exposed under UV for 30 min before the experiment to further eliminate the possible contaminations. For each single cell, the operation procedure on-chip is shown in Figure 1b. We washed mESCs with phosphate-buffered saline (PBS) and suspended the cells to a concentration of about 10^4–10^5 cells/mL and then used compress air to drive the cell suspension into the common channel of a microfluidic device. After the common channel is filled, we tuned the inlet pressure to a low level. We continuously monitored the cell loading procedure under the microscope to ensure that each reaction chamber has only one cell inside. When a single cell flew into an amplification unit, we manually actuated the corresponding valves to trap the cell and isolate it from others. We repeated this procedure for loading other single cells. Usually, it took less than 5 min to load 6 single cells and 2 negative controls into the chambers for a single chip. We rinsed the common channel with PBS to remove untrapped cells or debris, flushed the channel with compressed air, then rinsed it with water, and finally flushed it with air.

The WGA process starts with cell lysis. We opened the trapping valves and transferred single cells to lysis chambers using lysis buffer (30 mM Tris-HCl, 10 mM KCl, 5 mM EDTA, 0.5% Triton-X100, with 2 mg/mL Protease (Qiagen, CA, USA)) that freshly filled in the common channel. The lysis step took 90 min at 50 °C. Then, the protease in the lysate was inactivated at 75 °C for 20 min. The common channel was then thoroughly rinsed with water and flushed with air at 4 °C. We then injected the MALBAC preamplification buffer into the reaction and filled the first two sections in each reaction pipeline. The mixing was completed by passive diffusion, and we then raised the temperature to 95 °C for 3 min, and then ran 10 cycles of MALBAC preamplification (for each cycle: 30 s at 20 °C, 30 s at 30 °C, 30 s at 40 °C, 30 s at 50 °C, 30 s at 60 °C, 180 s at 70 °C, 20 s at 95 °C, and 10 s at 58 °C). The chip was held at 4 °C when the cycles were done, and the common channel was cleaned by water and air. Finally, the MALBAC PCR buffer was loaded into the chambers. The reagents were flushed into the chamber and mixed with the preamplification product through diffusion. We first held the chip at 95 °C for 3 min and then run 16 cycles of MALBAC PCR amplification (for each cycle: 20 s at 95 °C, 10 s at 58 °C, and 180 s at 70 °C). We collected the amplification product using pipet tips from outlets of the chambers.

We built a customized thermocycler coupled with an imaging system for the microfluidics WGA devices (Figure 1c). We put two Peltier devices between a water-cooled heat sink and an aluminum block. We bonded the polydimethylsiloxane (PDMS) chip with a silicon wafer that has intrinsically good heat conductance. A telecentric lens (2.0× Silver Series, Edmund Optics, NJ, USA) with an illuminator was set above the chip. We monitored the loading, lysis, and amplification steps of every single cell in the experiments using a monocolor CMOS camera (EO-5012M, 5 mega pixels, 1/2 in. size, Edmund Optics, NJ, USA). The cell could be clearly observed through light scattering without specific labeling (Figure 1d).

Each single-cell WGA reaction yielded a total amount of DNA around 50 ng, which reflects ca. 8000-fold amplification from the genomic DNA in one single cell. The amplification was reproducible and validated by quantitative PCR (qPCR, Table S1 and Figure S1, Supporting Information). We randomly picked six single-cell amplification products to construct libraries for next-generation sequencing using the Illumina HiSeq platform. In addition, two single mESCs were selected to perform the MALBAC reactions in tube for

dx.doi.org/10.1021/ac5032176 | Anal. Chem. 2014, 86, 9386–9390

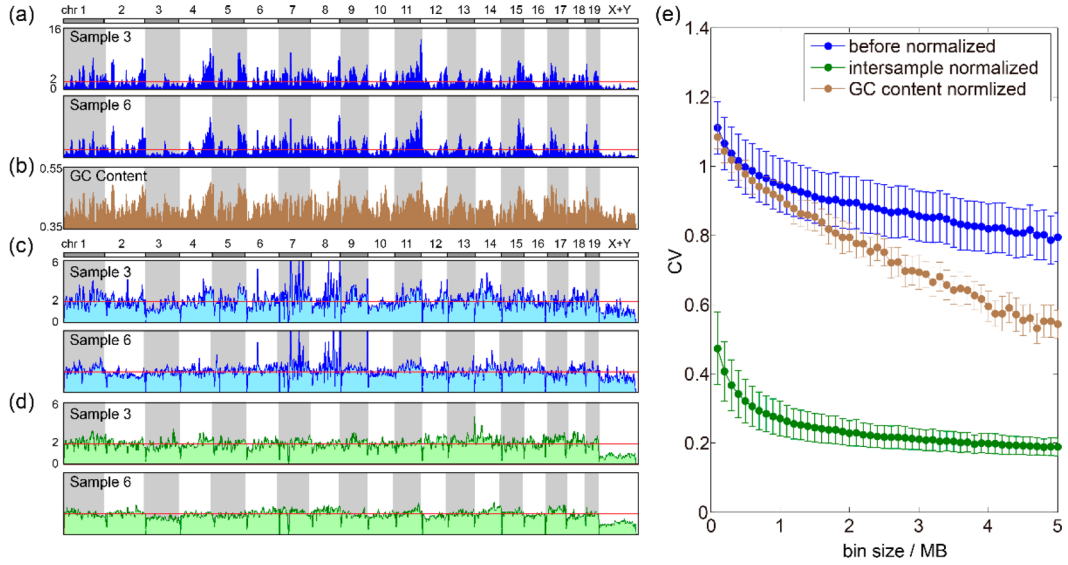


Figure 2. Sequencing coverage ratios of single mouse ES cells. Sequencing coverage depth (a) and GC content (b) across the whole genome of two single cells from two separate chips. (c) GC content normalization of samples #3 and #6, which exhibit effective reduction of the amplification bias. (d) Intersample normalized sequencing coverage ratios of samples #3 and #6. The binning window is 3 Mb. The red line indicates a ratio of 2. (e) The CV of coverage fluctuation of the four normal cells before and after the correction, with both GC content or cross-sample normalization approaches, in different binning window sizes.

comparison. We sequenced about 0.7 G bases for each library with 100bp paired-end reads.

We first estimated the contamination level of all the samples by analyzing the species origin of the sequencing data. The sequencing data was uploaded to Metagenomics-RAST[23] and processed using a standard routine. We obtained a group of sequences featuring a hit to the metagenome database and sorted them in taxonomic hit distributions. We removed the sequence annotated to Rodentia. We also removed the unassigned or unclassified sequences. In each library, the contamination level was presented by dividing the number of non-Rodentia-originated sequences by the total number of sequences that passed the quality control. The WGA on the microfluidic device showed a significant drop in contamination level (Figures S2 and S3, Supporting Information). For the libraries prepared through conventional MALBAC reaction in-tube, the average contamination level was about 4.8%, while those prepared through the microfluidic approach had a contamination level reduced by half. The average contamination ratios of libraries produced from microfluidic chip 1 (3 samples) and microfluidic chip 2 (3 samples) were 2.2% and 2.6%, respectively.

Figure 2a shows the distributions of sequencing coverage ratio across the whole genome for 2 of the 6 single mESCs from 2 chips (also see Table S2, Supporting Information). For each plot, the coverage ratio has been rescaled by the mean read-count of each sample. The plots clearly show the similar sequencing coverage patterns between different individual cells, indicating that MALBAC has intrinsic sequence-dependent amplification bias. Uneven amplification across the whole genome will introduce difficulties to identify the copy number variations (CNVs), which are the most important signatures for cancers and other diseases.[24−26] However, the sequence-dependent amplification bias provides a unique opportunity to perform the normalization of the sequencing coverage.

We first investigated the relationship between the amplification bias and the local GC content at the different genome sites. We calculated GC content across the whole mouse genome with a binning window size of 3 Mb. We compared the GC content pattern with the sequencing coverage depth across the whole genome and observed a strong positive correlation between them at 3 Mb resolution (Figure 2a,b). We then tested if the amplification bias can be simply corrected by just taking GC content into account. For each bin, the factor for CG content correction was calculated as below:

$$F = (C - 0.35)^2$$

where F was the correction factor and C was the GC content in the selected bin. The sequencing coverage ratio in the selected bin was divided by the correction factor and finally normalized by the average of the corrected ratios across the whole genome. The normalization exhibited effective reduction of the amplification bias (Figure 2c). However, notable fluctuations of coverage could still be found in some regions, especially in those with low GC content.

We then tested if the rectification procedure could be further improved by normalizing the sequencing coverage through the intersample correction, using the coverage curve of one representative cell as the standard. We first performed the cross-normalization between any of the two cells from the six sequenced samples using large binning windows (10 and 50 Mb) and calculated the coefficient of variation (CV) of each correction result (Tables S3 and S4, Supporting Information). We found that, as expected, not every single cell was identical and the cross-sample normalization unveiled the noticeable difference between samples. We thus chose the average coverage curve of three samples with lower average CV values (samples 2, 5, 6 for sample 3, and samples 2, 3, 5 for sample 6) as the "standard" curve to perform the normalization for all samples (Figure 2d). Samples 2, 3, 5, and 6 showed normal copy numbers, two for all the autosomes, along the whole genome.

We compared the CV of the coverage fluctuation of the four normal cells before normalization with that after two normalization methods, GC content correction and cross-

dx.doi.org/10.1021/ac5032176 | Anal. Chem. 2014, 86, 9386−9390

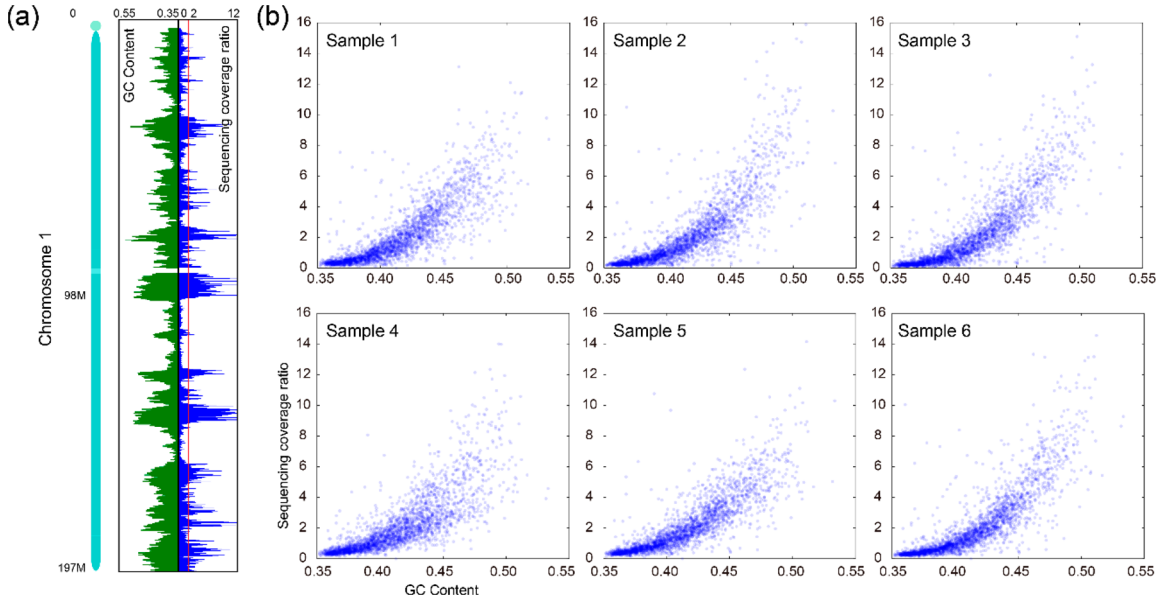


Figure 3. (a) The correlation between GC content and sequencing coverage depth across the chromosome 1. Binning window size is 100 kb. (b) Correlation between GC content and sequencing coverage depth of six single-cell samples. Binning window size is 1 Mb.

sample correction, using different sizes of binning windows (Figure 2e). Both the GC content correction and intersample normalization process could significantly improve the accuracy and resolution of copy number identification. Among these two correction approaches, the intersample normalization exhibited lower noise and CV, while a pure GC content correction demonstrated the great potential to identify the CNVs without "standard sample".

Each MALBAC primer contains eight random nucleotides, in which part the GC ratio is measured to be 78%. We found that, in our microfluidic MALBAC experiments, the amplification product exhibited a GC ratio of 50% (Table S5, Supporting Information), which was substantially higher than that of 40% for the entire mouse genome. This observation implies that MALBAC amplification favors GC rich regions during the reaction,[17] probably due to the GC dominant random section of the primers. Furthermore, this correlation can still be observed when the binning window is reduced to 100kb (Figure 3a); higher GC content greatly facilitated the amplification and consequently elevated the coverage (Figure 3b), causing the amplification bias across the whole genome.

The MALBAC reaction on the microfluidic platform has several advantages over the conventional operations performed in micro PCR-tubes. First, the small reaction volume (total 1.1 μL per reaction) significantly reduces the cost per sample to 1/30 of the tube-based MALBAC reactions, making the potential large-scale single-cell sequencing survey of the complex system more affordable. Furthermore, the reduction of cell lysis volume and reaction volume also significantly eliminates the exogenous contamination that is inevitably embedded in the reagents. Small and enclosed microfluidic environments also make the reactions more robust against contamination.[27] Third, the cell selection and confirmation can be continuously monitored under a microscope, ensuring that only single cells are isolated into the reaction chambers and the cells are completely lysed. Fourth, with lithographically defined reaction chambers, the multistep amplification reactions will be more reproducible than conventional tube-based operations.[20] Last but not least, the microfluidic platform enables great flexibility of integration with other functions such as cell staining, sorting,[28−30] and enrichment[31] at the front end, as well as the potential to extend the throughput by simply adding more reaction pipelines on one single chip.

The sequence-dependent amplification efficiency enables the intersample correction of sequencing coverage bias, leading to a more accurate identification of copy number in a higher resolution. A previous study using a single *E. coli* cell shows that MALBAC has inherent bias in amplification which is not strongly affected by gain.[27] Our result on mammalian cells suggests that the intrinsic amplification bias probably comes from GC content of different sequence regions, which further affect the annealing and extending efficiency during the multiple preamplification cycles. By a simple correction factor calculated from GC content, the sequencing coverage ratio can be corrected without a standard sample. Although the result of this method shows higher noise and CV, this approach provides an *ab initio* method to improve the accuracy of CNV determination for those applications in which no "standard" cells exist due to the intrinsic heterogeneity between cells, for example, the single cells isolated from cancer tissues. We also envision that by better fitting the sequencing coverage ratio−GC content curve, a more detailed procedure may be further optimized to improve the correction.

In summary, we present a novel microfluidic device to perform the high-throughput single-cell MALBAC whole-genome amplification reactions. This device greatly reduces the cost of reaction for each sample and eliminates the exogenous contamination during WGA. The whole-genome amplification reactions are reproducible, enabling the robust sequencing coverage correction between the samples to achieve accurate identification of copy number variations across the whole genome at a fine scale. We also discover the relationship between GC content of genome region and the amplification bias with our methods. The *ab initio* correction through GC content can noticeably improve the amplification evenness of MALBAC without using any "standard" cell as a reference. Our microfluidic method is suitable for whole-genome amplification of multiple single cell samples with a minimum requirement of

dx.doi.org/10.1021/ac5032176 | *Anal. Chem.* 2014, 86, 9386−9390

single-cell handling skills. In the further development, the device can be extended to a higher throughput and the technique may be integrated with other microfluidic cell analysis technologies.

■ ASSOCIATED CONTENT

Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: yanyi@pku.edu.cn.

Notes

The authors declare the following competing financial interest(s): S.L. is the founder and shareholder of Yikon Genomics..

■ ACKNOWLEDGMENTS

The authors thank Xinglong Wu, Xiaoying Fan, and Prof. Fuchou Tang for providing samples, Haiwei Qiu and Zitian Chen for help and discussion on building the thermocycler, and Chunmei Li for preparing the libraries. This work was supported by the National Natural Science Foundation of China (21222501, 21327808, and 91313302 to Y.H.)

■ REFERENCES

(1) Blainey, P. C.; Quake, S. R. *Nat. Methods* **2014**, *11*, 19−21.

(2) Wang, J.; Fan, H. C.; Behr, B.; Quake, S. R. *Cell* **2012**, *150*, 402−412.

(3) Navin, N.; Kendall, J.; Troge, J.; Andrews, P.; Rodgers, L.; McIndoo, J.; Cook, K.; Stepansky, A.; Levy, D.; Esposito, D.; Muthuswamy, L.; Krasnitz, A.; McCombie, W. R.; Hicks, J.; Wigler, M. *Nature* **2011**, *472*, 90−94.

(4) Lu, S.; Zong, C.; Fan, W.; Yang, M.; Li, J.; Chapman, A. R.; Zhu, P.; Hu, X.; Xu, L.; Yan, L.; Bai, F.; Qiao, J.; Tang, F.; Li, R.; Xie, X. S. *Science* **2012**, *338*, 1627−1630.

(5) Xue, Z.; Huang, K.; Cai, C.; Cai, L.; Jiang, C.-y.; Feng, Y.; Liu, Z.; Zeng, Q.; Cheng, L.; Sun, Y. E.; Liu, J.-y.; Horvath, S.; Fan, G. *Nature* **2013**, *500*, 593−597.

(6) Yan, L.; Yang, M.; Guo, H.; Yang, L.; Wu, J.; Li, R.; Liu, P.; Lian, Y.; Zheng, X.; Yan, J.; Huang, J.; Li, M.; Wu, X.; Wen, L.; Lao, K.; Li, R.; Qiao, J.; Tang, F. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1131−1139.

(7) Hou, Y.; Fan, W.; Yan, L.; Li, R.; Lian, Y.; Huang, J.; Li, J.; Xu, L.; Tang, F.; Xie, X. S.; Qiao, J. *Cell* **2013**, *155*, 1492−1506.

(8) Ramskold, D.; Luo, S.; Wang, Y.-C.; Li, R.; Deng, Q.; Faridani, O. R.; Daniels, G. A.; Khrebtukova, I.; Loring, J. F.; Laurent, L. C.; Schroth, G. P.; Sandberg, R. *Nat. Biotechnol.* **2012**, *30*, 777−782.

(9) Ni, X.; Zhuo, M.; Su, Z.; Duan, J.; Gao, Y.; Wang, Z.; Zong, C.; Bai, H.; Chapman, A. R.; Zhao, J.; Xu, L.; An, T.; Ma, Q.; Wang, Y.; Wu, M.; Sun, Y.; Wang, S.; Li, Z.; Yang, X.; Yong, J.; Su, X.-D.; Lu, Y.; Bai, F.; Xie, X. S.; Wang, J. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 21083−21088.

(10) Novarino, G.; Fenstermaker, A. G.; Zaki, M. S.; Hofree, M.; Silhavy, J. L.; Heiberg, A. D.; Abdellateef, M.; Rosti, B.; Scott, E.; Mansour, L.; Masri, A.; Kayserili, H.; Al-Aama, J. Y.; Abdel-Salam, G. M. H.; Karminejad, A.; Kara, M.; Kara, B.; Bozorgmehri, B.; Ben-Omran, T.; Mojahedi, F.; Mahmoud, I. G. E. D.; Bouslam, N.; Bouhouche, A.; Benomar, A.; Hanein, S.; Raymond, L.; Forlani, S.; Mascaro, M.; Selim, L.; Shehata, N.; Al-Allawi, N.; Bindu, P. S.; Azam, M.; Gunel, M.; Caglayan, A.; Bilguvar, K.; Tolun, A.; Issa, M. Y.; Schroth, J.; Spencer, E. G.; Rosti, R. O.; Akizu, N.; Vaux, K. K.; Johansen, A.; Koh, A. A.; Megahed, H.; Durr, A.; Brice, A.; Stevanin, G.; Gabriel, S. B.; Ideker, T.; Gleeson, J. G. *Science* **2014**, *343*, 506−511.

(11) Telenius, H. k.; Carter, N. P.; Bebb, C. E.; Nordenskjöld, M.; Ponder, B. A. J.; Tunnacliffe, A. *Genomics* **1992**, *13*, 718−725.

(12) Cheung, V. G.; Nelson, S. F. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14676−14679.

(13) Dean, F. B.; Nelson, J. R.; Giesler, T. L.; Lasken, R. S. *Genome Res.* **2001**, *11*, 1095−1099.

(14) Dean, F. B.; Hosono, S.; Fang, L.; Wu, X.; Faruqi, A. F.; Bray-Ward, P.; Sun, Z.; Zong, Q.; Du, Y.; Du, J.; Driscoll, M.; Song, W.; Kingsmore, S. F.; Egholm, M.; Lasken, R. S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5261−5266.

(15) Paez, J. G.; Lin, M.; Beroukhim, R.; Lee, J. C.; Zhao, X.; Richter, D. J.; Gabriel, S.; Herman, P.; Sasaki, H.; Altshuler, D.; Li, C.; Meyerson, M.; Sellers, W. R. *Nucleic Acids Res.* **2004**, *32*, No. e71.

(16) Pugh, T. J.; Delaney, A. D.; Farnoud, N.; Flibotte, S.; Griffith, M.; Li, H. I.; Qian, H.; Farinha, P.; Gascoyne, R. D.; Marra, M. A. *Nucleic Acids Res.* **2008**, *36*, No. e80.

(17) Zong, C.; Lu, S.; Chapman, A. R.; Xie, X. S. *Science* **2012**, *338*, 1622−1626.

(18) Marcus, J. S.; Anderson, W. F.; Quake, S. R. *Anal. Chem.* **2006**, *78*, 956−958.

(19) White, A. K.; VanInsberghe, M.; Petriv, O. I.; Hamidi, M.; Sikorski, D.; Marra, M. A.; Piret, J.; Aparicio, S.; Hansen, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 13999−14004.

(20) Streets, A. M.; Zhang, X.; Cao, C.; Pang, Y.; Wu, X.; Xiong, L.; Yang, L.; Fu, Y.; Zhao, L.; Tang, F.; Huang, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 7048−7053.

(21) Marcy, Y.; Ishoey, T.; Lasken, R. S.; Stockwell, T. B.; Walenz, B. P.; Halpern, A. L.; Beeson, K. Y.; Goldberg, S. M. D.; Quake, S. R. *PLoS Genet.* **2007**, *3*, e155.

(22) Unger, M. A.; Chou, H.-P.; Thorsen, T.; Scherer, A.; Quake, S. R. *Science* **2000**, *288*, 113−116.

(23) Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A.; Wilkening, J.; Edwards, R. *BMC Bioinf.* **2008**, *9*, 386.

(24) Pollack, J. R.; Perou, C. M.; Alizadeh, A. A.; Eisen, M. B.; Pergamenschikov, A.; Williams, C. F.; Jeffrey, S. S.; Botstein, D.; Brown, P. O. *Nat. Genet.* **1999**, *23*, 41−46.

(25) Shlien, A.; Malkin, D. *Genome Med.* **2009**, *1*, 62.

(26) Zhang, F.; Gu, W.; Hurles, M. E.; Lupski, J. R. *Annu. Rev. Genomics Hum. Genet.* **2009**, *10*, 451−481.

(27) de Bourcy, C. F. A.; De Vlaminck, I.; Kanbar, J. N.; Wang, J.; Gawad, C.; Quake, S. R. *PLoS One* **2014**, *9*, No. e105585.

(28) Kovac, J. R.; Voldman, J. *Anal. Chem.* **2007**, *79*, 9321−9330.

(29) Shirasaki, Y.; Tanaka, J.; Makazu, H.; Tashiro, K.; Shoji, S.; Tsukita, S.; Funatsu, T. *Anal. Chem.* **2005**, *78*, 695−701.

(30) Sun, T.; Kovac, J.; Voldman, J. *Anal. Chem.* **2013**, *86*, 977−981.

(31) Lee, M. G.; Shin, J. H.; Bae, C. Y.; Choi, S.; Park, J.-K. *Anal. Chem.* **2013**, *85*, 6213−6218.

Supporting Information

A microfluidic whole genome amplification device for single cell sequencing

Zhilong Yu[1], Sijia Lu[2], and Yanyi Huang[1,*]

[1] Biodynamic Optical Imaging Center, and College of Engineering, Peking University, Beijing 100871, China
[2] Yikon Genomics Co. Ltd., 1 China Medical Ave., Taizhou, Jiangsu 225300, China

* To whom correspondence should be addressed. Email: yanyi@pku.edu.cn

1. Experimental

1.1 Mold fabrication

The silicon molds were fabricated by photolithography. We designed photomasks with CAD software (AutoCAD, Autodesk Inc., CA, USA). We used SU-8 2025 photoresist (MicroChem, MA, USA) to fabricate the mold of the control layer, with thickness of 40 μm by spinning the wafer at 2000 rpm for 60 s. The wafer was prebaked at 65 °C for 5 min and 95 °C for 10 min. The exposure took 10 s, then the wafer was postbaked at 65 °C for 5 min and 95 °C for 10 min. After development, the wafer was hard-baked at 150 °C for 2 h. We used AZ50XT photoresist (AZ Electronic Materials, NJ, USA) to fabricate the mold of the flow layer, with thickness of 50 μm by spinning the wafer at 1000 rpm for 60 s. The wafer was also prebaked at 65 °C for 5 min and 95 °C for 10 min. The exposure took 20 s and the mold was developed. The channels were rounded by thermal reflow.

1.2 PDMS chip fabrication

All devices were fabricated using multilayer soft lithography. The molds were first treated with trimethyl chlorosilane vapor for 5 min to prevent PDMS from bonding to the patterned photoresist. The flow layer was made from a mixture of PDMS (RTV615, A : B = 5:1, General Electric). 30 g mixture was poured on the mold, degased and baked for 18 min at 80 °C. The control layer was made from another mixture of PDMS (RTV615, A : B = 20:1). The mixture of PDMS was coated on the control mold by spinning the mold at 1700 rpm for 1 min. Then the control layer was baked for 25 min at 80 °C. The two layers were manually aligned and boned together by baking at 80 °C for 40 min. A mixture of PDMS (RTV615, A : B = 10:1) was poured on a clean silicon wafer. The wafer was spinning at 900 rpm for 1 min and then baked at 80 °C for 7min. The chip was bonded to this substrate by baking at 80 °C over night. The chips were stored in clean room before use.

1.3 Reagents

MALBAC reaction reagents were purchased from Yikon Genomics (Taizhou, China). Reactions are supplemented with additional deoxy-ribonucleotide triphosphates (0.3 mM) and primer (0.5 μM). PCR grade water was purchased from Ambion (Life Technologies, NY, USA).

1.4 Primer sequences

The MALBAC primers and quantitative PCR primers were purchased from Invitrogen (China). The MALBAC pre-amplification primers have a common 27-nucleotide sequence and 8 variable nucleotides while the MALBAC PCR amplification primer only has the common sequence. The common 27-nucleotide sequence is GTG AGT GAT GGT TGA GGT AGT GTG GAG. Primer sequences for quantitative PCR to test the amplification results of single-cell whole genome amplification are listed in Supplementary Table S1.

1.5 Cell loading procedure

mESCs was washed and suspended to phosphate-buffered saline (PBS) for loading. The whole procedure was monitored under the home-built imaging system. Before the operation, all the valves were closed. The cell suspension was first loaded into a piece of silicone micro-tubing. Then the control valve of the common channel inlet was opened. The cell suspension was driven into the common channel using compress air (~15 psi). After the common channel is filled, we tuned the inlet pressure to a low level (~1 psi). By quickly opening and closing the valve of the common channel outlet, a target cell in the common channel can be moved to the inlet of a target chamber. The inlet valve of the target chamber was opened and the cell was flushed into the selected amplification unit. After the cell was loaded into the chamber, the inlet valve of the selected chamber was closed again. We repeated this procedure for loading other single cells into other chambers. Finally the common channel was rinsed with PBS to remove untrapped cells or debris, and flushed with compressed air, and then rinsed with water, and finally flushed with air again. For a single chip with six single cells and two negative controls, it usually takes less than 5 min to complete the cell loading procedure.

2. Supplementary Tables

Supplementary Table S1: Primers for quantitative PCR to test the amplification results of single cell whole genome amplification.

Primer	Sequence
Chr1	CAGCCAGGTGAAGACTTTCTTTGTA
	GAGAGCTATTTTTGCAGTTCCATCA
Chr2	CATGACCTGTACACCACTTCATCAG
	ATCTGCCTATTTGCTTAGGATGGAG
Chr3	GAAGCGAATCACAGTAGAGAACAGC
	CCCTTACAGGCTGAGTTCTTCAAAT
Chr4	GATTCCTCAGTTATGTTGGCAGAGA
	GAACAAGCAAGCTTTGAGAGTTGAC
Chr5	GTTCTGTGTGTGTGAAATCAGGATG
	CAATATGCATCTTATCTGGGCTGAC

Supplementary Table S2: Quantitative PCR results of 12 single cell samples and 4 negative controls (NC). C_t (threshold cycle) values of different primers are shown.

Chip	Sample	Chr1 C_t	Chr2 C_t	Chr3 C_t	Chr4 C_t	Chr5 C_t	Amplified	Selected as
1	1	21.5	24.4	N/A	21.9	28.5	Yes	
1	2	30.1	22.7	25.2	26.8	22.1	Yes	Sample 1
1	3	22.0	25.8	26.9	23.2	22.4	Yes	
1	4	24.2	24.6	22.0	21.7	27.0	Yes	Sample 2
1	5	24.4	22.9	22.3	20.5	27.7	Yes	
1	6	19.0	24.6	31.0	21.7	29.4	Yes	Sample 3
1	NC1	N/A	27.2	N/A	N/A	N/A	No	
1	NC2	34.0	33.5	33.5	N/A	N/A	No	
2	1	23.7	31.3	33.6	N/A	29.7	No	
2	2	21.9	35.7	22.1	29.5	31.2	Yes	Sample 4
2	3	25.5	21.1	26.6	20.8	25.3	Yes	Sample 5
2	4	25.1	24.2	20.6	22.8	24.7	Yes	
2	5	22.8	26.6	22.4	N/A	26.0	Yes	
2	6	23.6	25.8	23.0	21.7	26.8	Yes	Sample 6
2	NC1	33.5	33.7	18.1	N/A	N/A	No	
2	NC2	36.6	33.3	N/A	N/A	37.3	No	

Supplementary Table S3: Average CV of cross-normalization between any of two cells from the six sequenced samples. Binning windows size is 10 Mb

Standard Sample	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
Sample 1	-	0.319	0.265	0.441	0.233	0.268
Sample 2	0.381	-	0.223	0.364	0.260	0.208
Sample 3	0.288	0.240	-	0.389	0.241	0.171
Sample 4	0.526	0.382	0.459	-	0.337	0.355
Sample 5	0.261	0.258	0.287	0.312	-	0.213
Sample 6	0.290	0.237	0.182	0.347	0.207	-
Average	0.349	0.287	0.283	0.371	0.256	0.243

Supplementary Table S4: Average CV of cross-normalization between any of two cells from the six sequenced samples. Binning windows size is 50 Mb

Standard Sample	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
Sample 1	-	0.287	0.234	0.443	0.223	0.261
Sample 2	0.344	-	0.168	0.319	0.184	0.169
Sample 3	0.245	0.173	-	0.352	0.173	0.130
Sample 4	0.536	0.371	0.439	-	0.331	0.354
Sample 5	0.242	0.168	0.188	0.302	-	0.176
Sample 6	0.274	0.173	0.130	0.321	0.158	-
Average	0.328	0.234	0.232	0.348	0.214	0.218

Supplementary Table S5: GC content in the random region of MALBAC primer and GC content in total reads from sequencing data.

Sample	GC content in random region of primer	GC content in total reads
Sample 1	78.5%	50.1%
Sample 2	77.5%	50.3%
Sample 3	78.9%	50.9%
Sample 4	77.8%	49.4%
Sample 5	77.8%	49.2%
Sample 6	78.8%	50.6%

3. Supplementary Figures

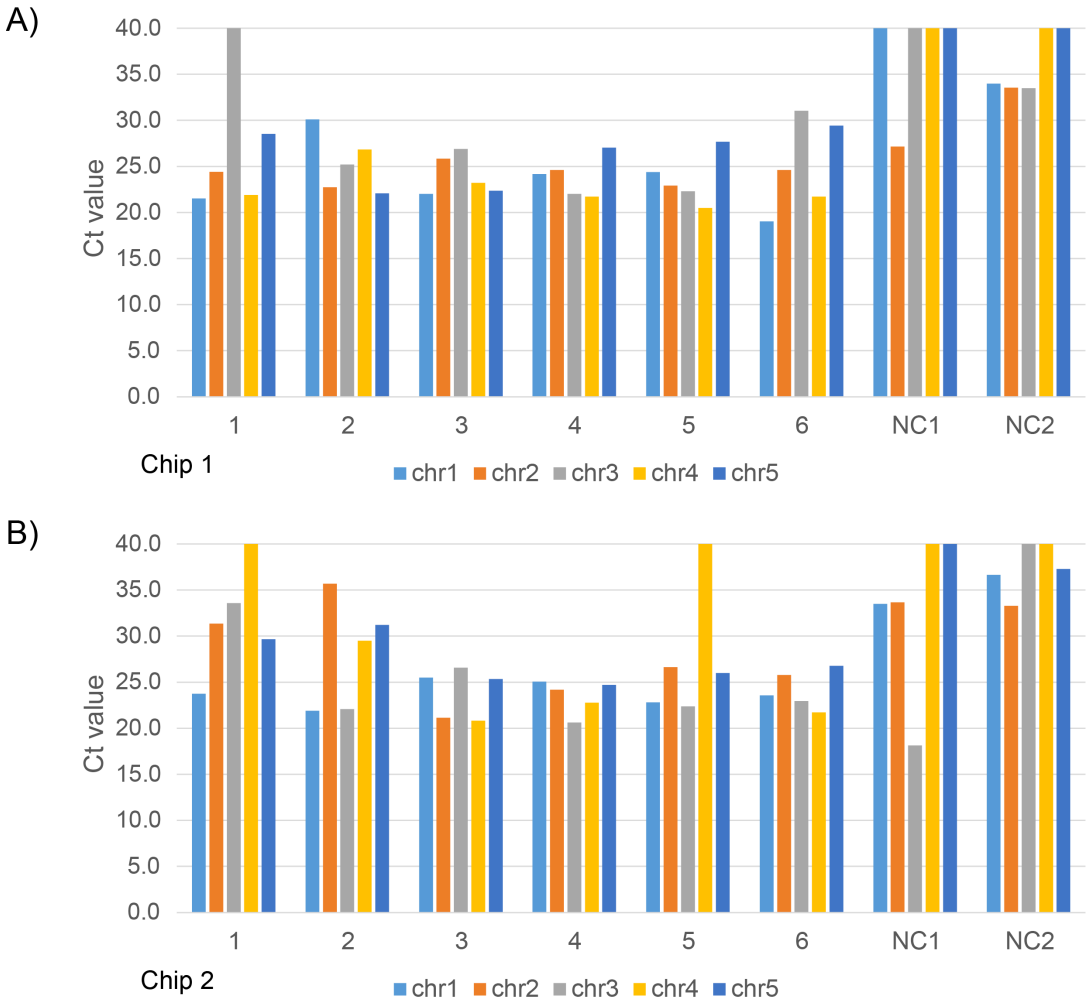


Figure S1. Quantitative PCR results of Chip 1 (A) and Chip 2 (B), including 12 single cell samples and 4 negative controls (NC).

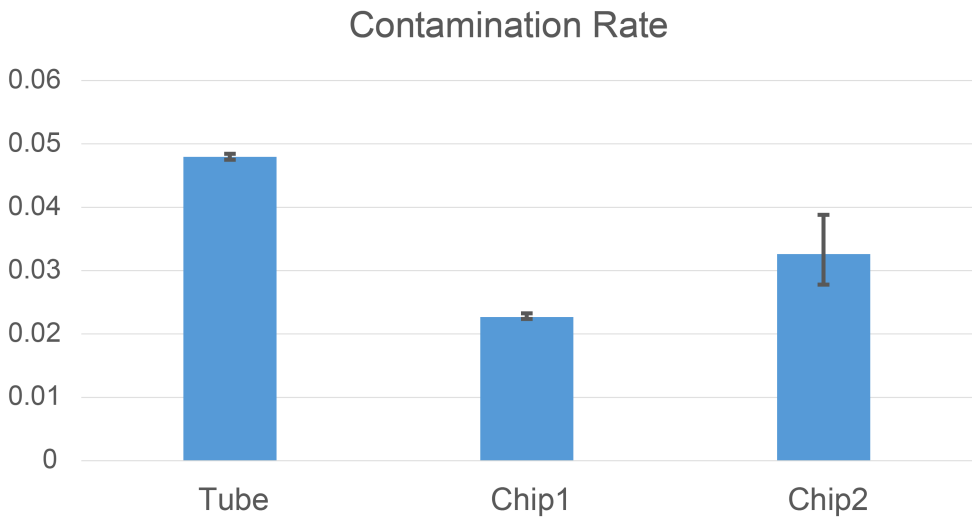


Figure S2. Comparison of contamination level of single cell MALBAC reactions between conventional tube method and microfluidic method.

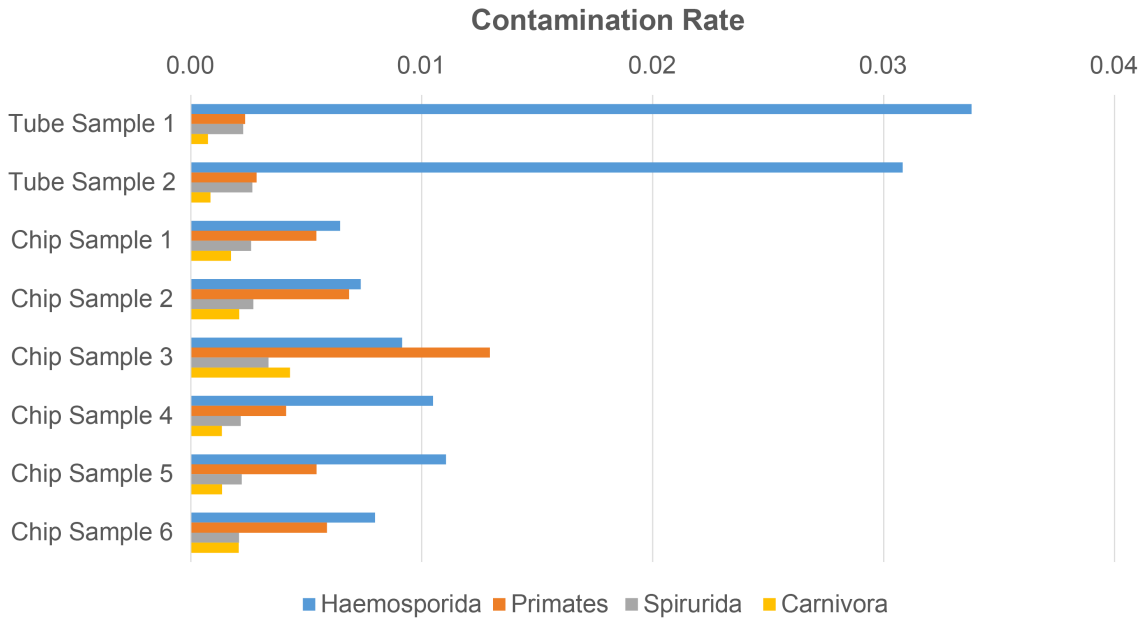


Figure S3. Top 4 contamination sources in tube samples and chip samples organized in order. Compared to tube samples, contamination from Haemosporida is significantly reduced in chip samples.