Electronic Supplementary Material (ESI) for Lab on a Chip.

Device for Whole Genome Sequencing Single Circulating Tumor Cell from Whole Blood

Ren Li[a,b,c,‡], Fei Jia[a,b,‡], Weikai Zhang[d], Fanghao Shi[a,b], Zhiguo Fang[a,b], Hong Zhao[e*], Zhiyuan Hu[a,b,f,g,*] and Zewen Wei[d*]

[a] CAS Key Laboratory of Standardization and Measurement for Nanotechnology, CAS Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, CAS Center for Excellence in Nanoscience, National Center for Nanoscience and Technology of China, Beijing 100190, China.

[b] University of Chinese Academy of Sciences, Beijing 100049, PR China

[c] Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, People's Republic of China

[d] Department of Biomedical Engineering, School of Life Science, Beijing Institute of Technology, Beijing, P. R. China

[e] Department of Hepatobiliary Surgery and Department of Pathology, State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China

[f] School of Nanoscience and Technology, Sino-Danish College, University of Chinese Academy of Sciences, Beijing 100049, P. R. China.

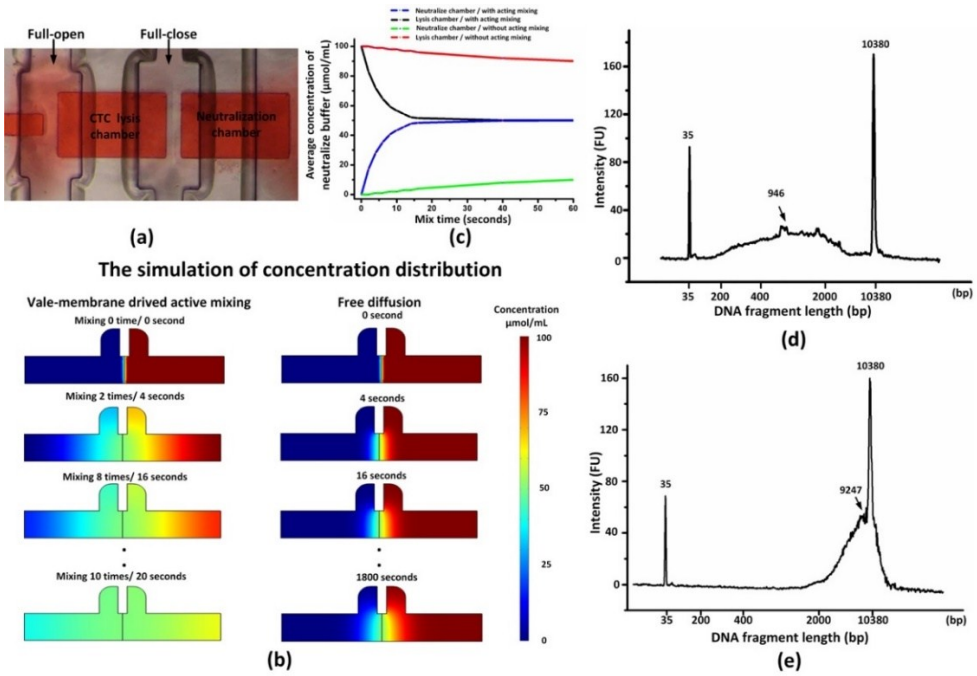
[g] Center for Neuroscience Research, School of Basic Medical Sciences, Fujian Medical University, Fuzhou 350108, Fujian Province, China.

‡ These authors contributed equally to the work.

* Corresponding Authors: Zewen Wei (weizewen@bit.edu.cn), Zhiyuan Hu (huzy@nanoctr.cn) or Hong Zhao (zhaohong@cicams.ac.cn)

Supplementary Data S1

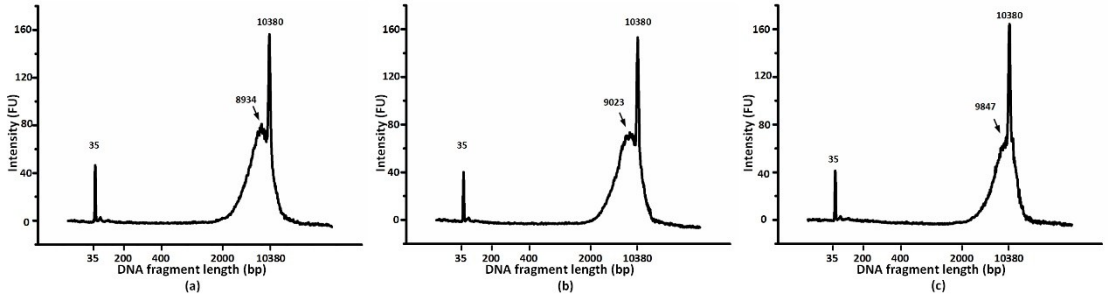
Tri-state valves as an active mixer



To mix different liquids in neighboring chambers, the membrane of tri-states valve was repetitively switched full-open and full-close (as shown in figure a). Using COMSOL Ver5.2 software, the membrane driven active mixing process was 2-D simulated (left images of figure (b)). The dimensional parameters were set according to the real microfluidic chip. The initial concentrations of liquid in neutralization (right) and lysis (left) chambers were set at 0 and 100 μmol/ml. Opening/closing each cycle of valve costs 2 s. As Control Group, a natural diffusion process was also simulated with the same parameters (right images of figure (b)). To ensure maximum diffusion, the valve was maintained full-open in the simulation regarding natural diffusion. Figure (c) shows the quantitative data of mean liquid concentration variation in lysis and neutralization chambers. As shown in the figure (c), active mixing group was far more insufficient under the natural diffusion mode. The MDA products of single cell, were examined using Agilent 2100 Electrophoresis Bioanalyzer (Agilent, USA). (d) The reagents were without tri-state valves mixing. The main peak of DNA fragment length reduced gradually (946 bp). The total DNA content and concentration also decreased. (e) The reagents were mixed 10 times on chip with the assistance of tri-state valves. The main peak of fragment length located in 9,247 bp, matching the typical value of MDA products, which should be around 10,000 bp. Moreover, the total DNA mass and concentration also satisfied the requirements of further PCR amplification and Sanger' sequencing.

Supplementary Data S2

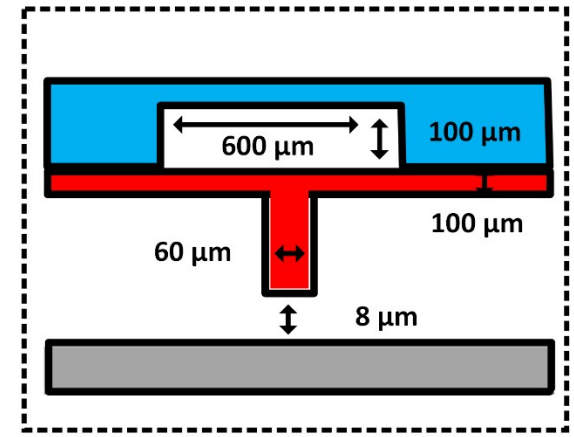
DNA fragments length of MDA products



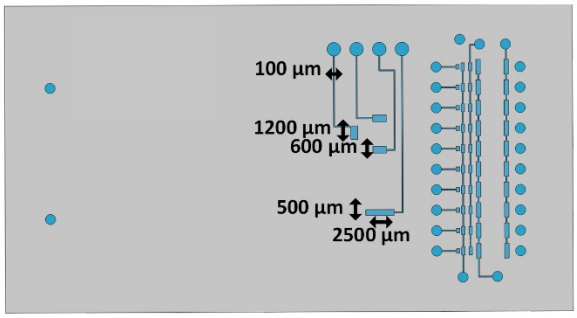
To ensure optimum whole-genome sequencing results, the MDA products were retrieved from our chip and thoroughly examined. We employed Qubit 3.0 fluorescence ration instrument (Thermo, USA) and Agilent 2100 bio-analyzer (Agilent, USA) to analyze MDA products and DNA fragments length. For 3 different amplified samples, the main peak values of fragments length were 8934, 9023 and 9847 bp, respectively. According to established standards, these parameters could satisfy the requirements for the whole genome sequencing.

Supplementary Data S3

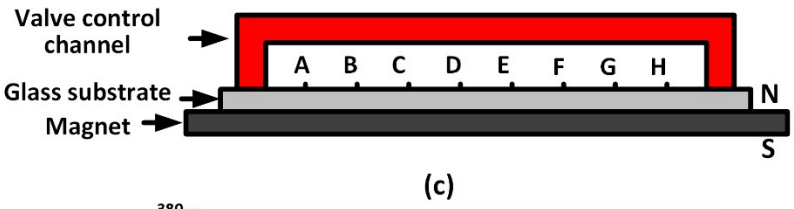
The dimensional parameters of microfluidic chip and magnetic field strength distribution



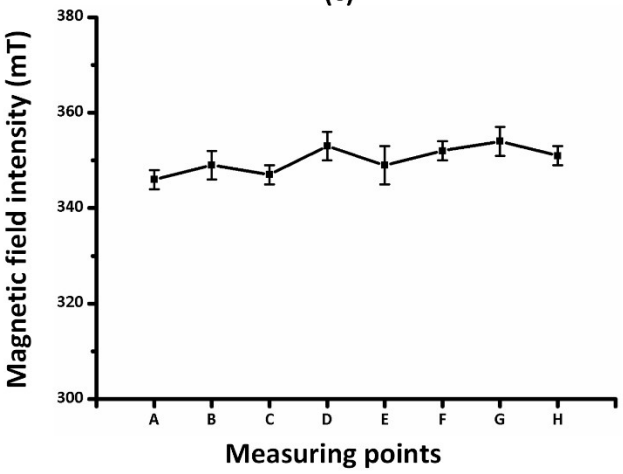
(a)



(b)



(c)

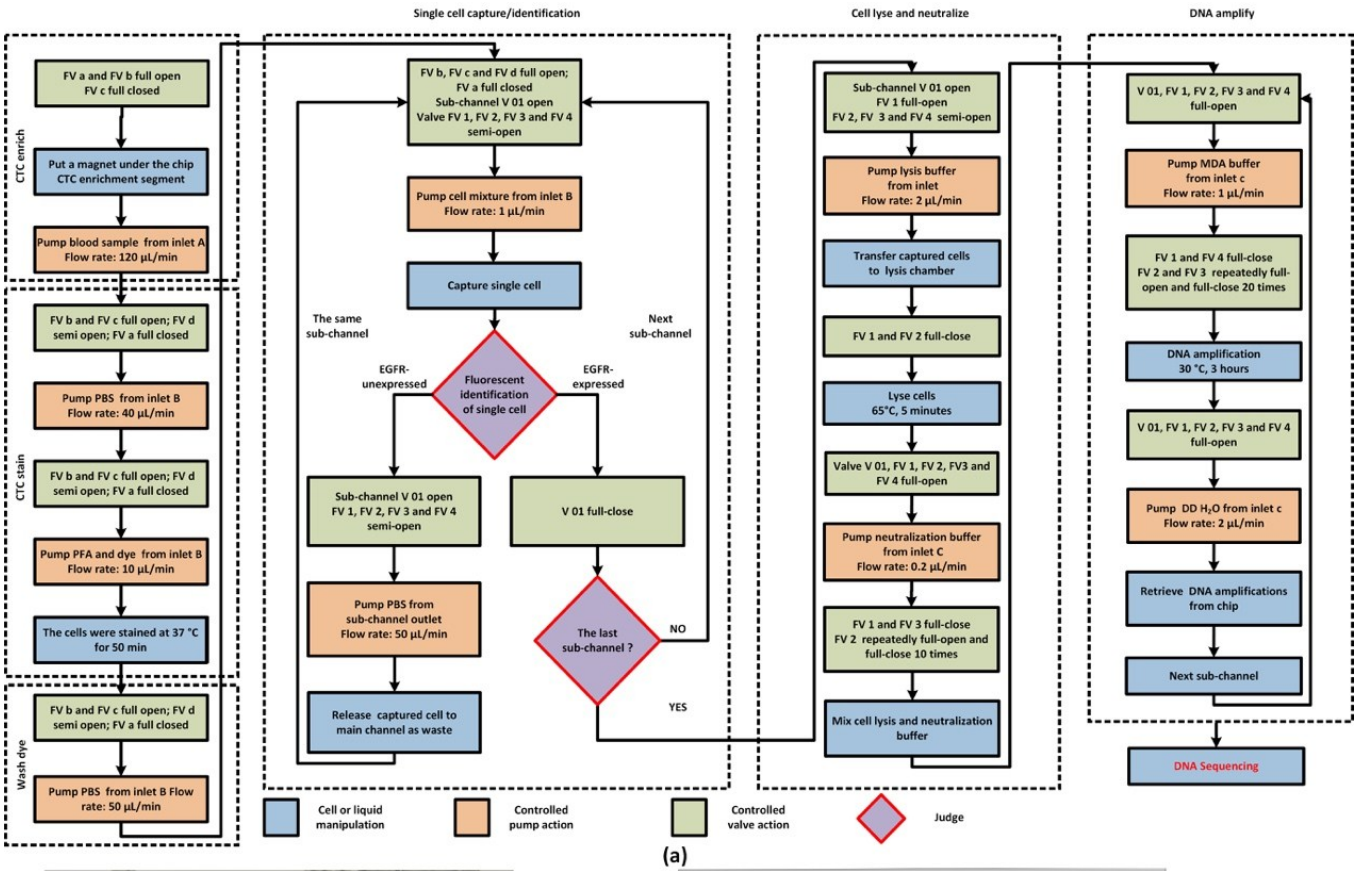


(d)

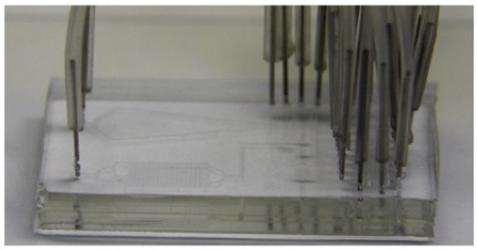
(a) and (b) show the dimensional parameters of tri-state valve and the SCIGA-Chip. To isolate EpCAM positive cells in the blood sample, an NbFeB permanent magnet (50 mm X 50 mm X 25 mm) was placed under the cell isolating segment (schemed as (c)) with N pole upwards and S pole downwards. We measured the magnetic field strength of 8 points, using a digital gaussmeter (HT20, Hengtong, China). (d) indicates that the magnetic field intensities in 8 points were evenly distributed, ranging from 340 mT to 360 mT.

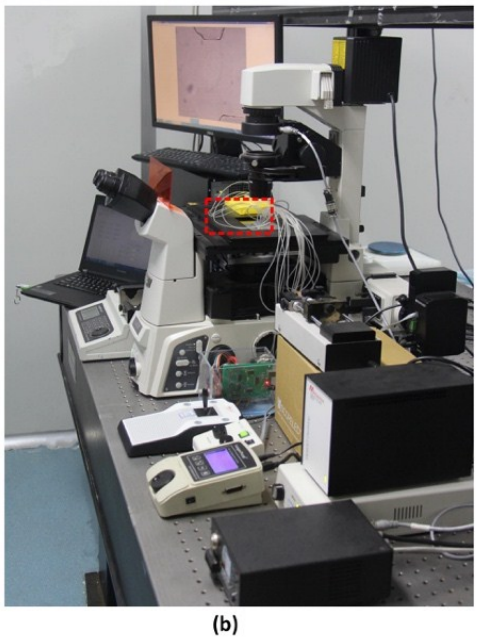
Supplementary Data S4

Automatic chip operation

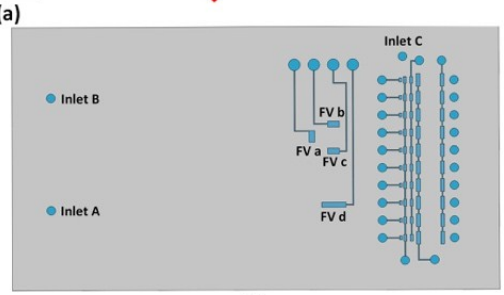


(a)

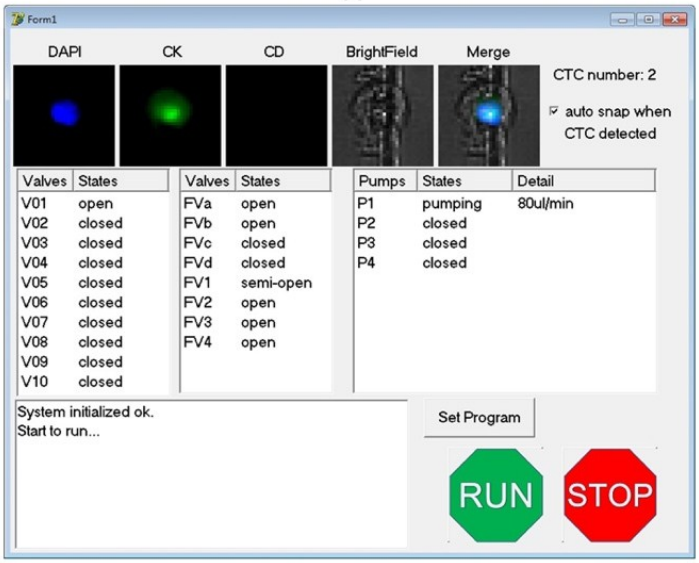




(b)



(c)



(d)

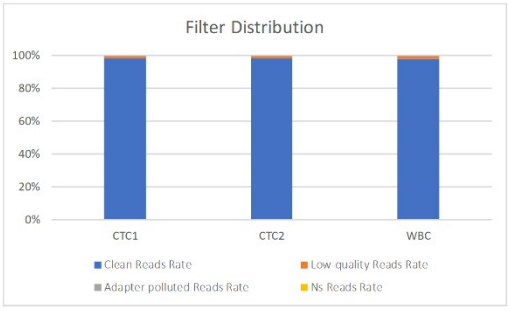
Figure (a) illustrates a logical flowchart of the controlling algorithm which consists of 6 main parts：CTC filtration and enrichment, CTC staining, cell washing, single cell capture & identification, cell lyse & neutralization and MDA. Figure (b) demonstrates the connection diagram of different off-chip controlling devices. Four syringe pumps, which were used to inject cell mixtures or reagents, were directly controlled by computer. An Intel 8051 MCU (Microcontroller Unit), which was controlled by computer, was used to switch the connection between air pressure / vacuum source and 18 solenoid valves. Hence, all on-chip valves were controlled by computer. Figure (b) is the photo of the system setup, which includes a Nikon fluorescent microscopy with CCD, a controlling computer, four syringe pumps, 18 solenoid valves and MCU. Figure (d) is the picture of the microfluidic chip in which only 10 valves No. 1 were connected for demonstration. Figure (c) shows the location of the inlet and the valves. Figure (d) shows the interface of the compiled algorithm running on Windows system.

Supplementary Data S5

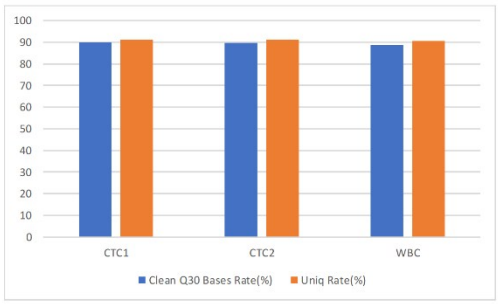
The quality of single CTC and WBC sequencing

Sample	Raw Reads Number	Raw Bases Number	Clean Reads Number	Clean Bases Number	Low-quality Reads Number	Adapter polluted Reads Number	Genome Length	Mapped Bases
CTC1	612676364	91901454600	603264296	90489644400	7625996	1786072	2897310462	67910769686
CTC2	625811068	93871660200	615696700	92354505000	8343498	1770870	2897310462	86688766285
WBC	621692074	93253811100	610169274	91525391100	9773438	1749362	2897310462	72178425297

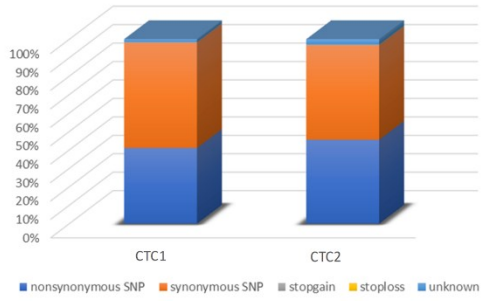
(a)



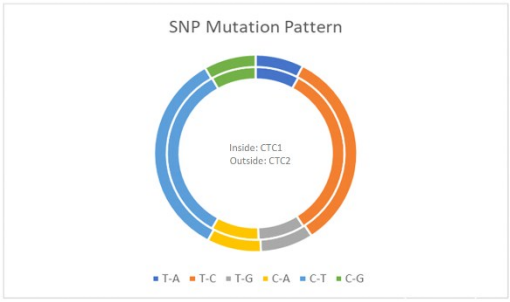
(b)



(c)



(d)



(e)

Figure (a) illustrates the fundamental statistics of whole genome sequencing data of CTCs and white blood cell, including raw/clean reads and bases number, low-quality/adapter polluted reads number, genome length and mapped bases. Figure (b) shows the distribution of the raw data. The clean reads rates in CTCs and WBC were greater than 98%. Besides, Low-quality reads, adapter polluted reads and Ns reads occupied the 2% of the distribution. Figure (c) demonstrates that Q30 bases rates in three parts were about 90% and the unique mapping rates of CTCs and WBC were more than 90%. Figure (d) shows the constitution of SNPs in CTCs. The synonymous SNP accounted for large parts and the meaningful part was nonsynonymous SNP, which was about 43% of all SNPs. Comparing with the mutation pattern in the Catalogue of Somatic Mutations in Cancer (COSMIC) Signatures of Mutational Processes in Human Cancer, mutations were dominated by C: G → T: A change, which were found in all cancer types and in most cancer samples. And T: A → C: G, which was another large part, was proved as a signature in liver cancer (figure (e)).