Fast, cost-effective and flexible DNA sequencing by roll-to-roll fluidics

Yanzhe Qin, Stephan A. Koehler, Yunyan Ling, Shiqiang Yu, Yongjie Zhang, Jie Luo, Kaijian Chen, Luo Junjie, Junjie Zeng, Haiyang Chu, Fei Wang, Wei Li, Dan Li, Xinran Yu, Xiangchao Wu, Shengming Zhao, Hao Lu, Ziqing Deng, Zhijian Yang, Ruibin Mai, Zhuo Liu, Zihua Niu, Xin Huang, Chengmei Xing, Xingcai Zhang, Yongyou Hu, Xinwen Peng, Xiaoxing Liao, Xiangmeng Qu, Yiming Gong, Qiushui Chen, Thomas M. Hermans, Wenwei Zhang & Huanghao Yang

Fast, cost-effective, and flexible DNA sequencing by roll-to-roll fluidics

Yanzhe Qin[1,2,3*], Stephan A. Koehler[4], Yunyan Ling[1], Shiqiang Yu[1], Yongjie Zhang[3], Jie Luo[1], Kaijian Chen[1], Junjie Luo[1], Junjie Zeng[1], Haiyang Chu[1], Fei Wang[1], Wei Li[1], Dan Li[1], Xinran Yu[5], Xiangchao Wu[1], Shengming Zhao[2], Hao Lu[2], Ziqing Deng[2], Zhijian Yang[3], Ruibin Mai[2], Zhuo Liu[2], Zihua Niu[2], Xin Huang[1], Chengmei Xing[1], Wenwei Zhang[1], Xun Xu[1], Xingcai Zhang[6], Yongyou Hu[7], Xinwen Peng[7], Yiming Gong[8], Qiushui Chen[3*] , Xiaoxing Liao[9], Xiangmeng Qu[10], Thomas M. Hermans[11*]& Huanghao Yang[3*]

[1]*BGI-Shenzhen, Shenzhen 518083, China.*

[2]*MGI, BGI-Shenzhen, Shenzhen 518083, China.*

[3]*New Cornerstone Science Laboratory, College of Chemistry, Fuzhou University, Fuzhou 350002, China.*

[4]*Artificial Intelligence and Machine Learning Laboratory, Riverside Research Institute, 70 Westview St Suite 3, Lexington, MA 02421, USA.*

[5]*Harvard School of Public Health, Boston, Massachusetts 02115, United States.*

[6]*Department of Materials Science and Engineering, Stanford University, Stanford, CA, 94305, USA*

[7]*School of Light Industry and Engineering, School of Environment and Energy, South China University of Technology, Guangzhou 510006, China.*

[8]*Department of Physics, Kyoto University, Kyoto 606-8502, Japan.*

[9]*Emergency and Disaster Medical Center,The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen 518107, China*

[10]*School of Biomedical Engineering, Sun Yat-Sen University, Shenzhen 518107, China*

[11]*IMDEA Nanoscience, C/ Faraday 9, Madrid 28049, Spain.*

[*]Corresponding author. E-mail: yanzheqin@126.com (Y.Q.); thomas.hermans@imdea.org (T.M.H.); hhyang@fzu.edu.cn (H.Y.); qchen@fzu.edu.cn (Q.C.)

Despite decreasing prices per DNA base pair over the past decade, small-scale next-generation sequencing (NGS) remains prohibitive in urgent clinical settings, since the reagent efficiency of sequence-by-synthesis (SBS) is low in the corresponding flow cells. Here we show NGS by roll-to-roll fluidics (r2r-fl), which is cost-effective for flexible biochip sizes. R2r-fl is a practical implementation of plane-Couette flow, with up to 85 times lower reagent consumption (US$ 0.16 per giga base pair), rinsing times of less than 2 seconds, and shortened turnaround times (TAT) for paired-end 100 bp (PE100) sequencing from days to under 12 hours. We maintain over 99.9% precision and 99.3% sensitivity of single nucleotide polymorphisms (SNP) for the human genome, 99.9% mapping rate for *E.coli*, and minimal nucleotide substitutions, deletions, or insertions in

the SARS-CoV-2 Alpha strain. Overall, r2r-fl makes NGS drastically faster and more affordable, thus contributing to revolutions in dealing with pathogens, cancers, genetic diseases, and personalized medicine.

Introduction

High-throughput DNA sequencing[1-3] is increasingly common in public health management, and has proven to be an effective tool for responding to epidemics, such as providing early warnings of regional outbreaks for SARS-CoV-2 or Ebola, [4, 5] as well as tracking of their evolution. [6-8] The core component of the current next-generation sequencing (NGS) system is the flow cell that was pioneered in 2005. [1] Typical cells have a 50-100 µm gap height with a glass slide covering a patterned biochip onto which single-stranded DNA molecules are adhered (see Extended Data Fig. 1a for commercially available cells).[3,9] To perform cycling SBS—involving DNA extension by a dNTP (deoxy-ribonucleoside triphosphate), fluorescence detection ('base calling') and regeneration[10,11] (see the Methods section), repetitively—the reagents need to be fully (99.9%) replaced to ensure extending by only a single complementary base in each cycle. To achieve that in a flow cell system, one needs to flush the cell chamber and tubing with a volume equal to several times their internal volume, due to the parabolic Poiseuille flow profile (Extended Data Fig. 1b-e). The pressure drop over the chamber $\Delta p \approx \frac{12\mu L V_a}{h^2}$, depends on the chamber length L [m], height h [m], dynamic viscosity μ [N·s/m] and the average fluid velocity V_a [m/s]. [12, 13] The maximum operating pressure for current sequencers is about ~90 kPa (~0.9 bar) due to pressure members such as manifolds and O-rings, which limits further performance improvement, such as: i) throughput, since this scales with chamber length L; ii) reagent efficiency, since for smaller gaps h less reagent is wasted; iii) turnaround time (TAT), since greater flow velocity V_a allows for faster reagent cycling.

Currently, the highest throughput flow cell in NGS platforms can generate 8 Tb at an average price of US$ 2 per giga base pair ($/Gb), with a TAT of two days. [14] However, for smaller biochips the reagent efficiency is getting lower, and prices surge. The high operational costs of NGS are due to fabrication

of flow cells with low tolerance, [15] and especially the use of expensive reagents (>90% of the total price, same for manufacturers). [16, 17] To lower reagent usage, Illumina launched the Nextseq 2000 and NovaSeq X, which recycles 30% of the reagents and triples the density of DNA molecule attachment points on the biochip, resulting in improved data throughput and enhanced reagent efficiency. Unfortunately, its potential is limited by the minimum size for the DNA attachment points to fulfill requisite brightness (~100 nm), and requiring super-resolution imaging as attachment points are closer than the Rayleigh limit which significantly prolongs the biochip scanning process. [18, 19] An alternative approach to reduce cost is to expand the biochip area, such as DNBSEQ-T7 and NovaSeq 6000. However, this approach may extend patient waiting time, as economies of scale are only possible when enough samples are collected and loaded. To reduce TAT, small flow cells such as Miseq are available that have a throughput of 4.5 Gb for PE150. However, the reagent usage efficiency is compromised due to the significant flushing requirement of the supply tubing, whose volume is much larger than the flow cell chamber itself (Table 1). Overall, flow cell platforms are subject to hydromechanical constraints imposed by Poiseuille flow, and engineers have to make a trade-off between cost, throughput and TAT. To overcome these constraints, BGI[20] and Ultima[21, 22] genomics have borrowed technologies from the semiconductor industry such as RCA cleaning or spin-coating in order to reduce the reagent thickness on the wafer. These arrangements get rid of the flow cell structure, using a naked wafer instead, and both companies have announced $100 human genome sequencing[23]. However, to be cost-effective both systems need to collect enough (patient) samples to process large wafer(s) before starting a single run, with a data production which far exceeds daily clinical demand (see Supporting Information section ‘Discussion on flow-cell free DNA sequencers’). Currently, there is a pressing clinical need for sequencers that can economically process small numbers of samples in a single run, which today’s technology cannot adequately meet.

Plane-Couette flow (pCF)—i.e., negligible pressure-drop shear-flow between two infinite large moving parallel plates—was first put forward by Reynolds in 1884, in describing lubrication theory. [24]

Experimental implementations, mainly to study high Reynolds number (*Re*) turbulent flow transitions, had to wait until 1954 by Reichardt [25]. In the few experimental records on pCF, a looped belt tensioned between rollers is fully immersed in fluid, [26,27] rendering it impractical for implementations other than fundamental fluid dynamics study. In contrast, circular Couette flow—a shear flow between two concentric cylinders—was rapidly adopted in the following years by other scientific communities, e.g., as a basis for rheology. [28,29] However, next we will show a practical non-immersed implementation of pCF we call roll-to-roll fluidics (r2r-fl, Fig. 1), which has very efficient mass transfer irrespective of biochip size, and is ideal for rapid and flexible scale NGS which meet clinical demand with low cost, high speed and flexible sample sizes.

Results

Establishment of Plane-Couette flow by roll-to-roll fluidics

Specifically, we combine slot-die injection to layer reagents at the bottom side of a moving PET (Polyethylene terephthalate) belt that grazes a stationary biochip (Fig. 1a). The flow behavior is schematically explained in the inset, where: i) upon slot die injection, ii) a plug flow layer of height $h/2$ is formed, which iii) expands and wets the biochip, forming plane-Couette flow of height h at average velocity V_a, and subsequently iv) is dragged by the PET belt after the biochip, reestablishing plug flow. Thus, the plug flow layer (Fig. 1a, *ii*) only needs to be half the height of that in the pCF zone (*iii*), due to conservation of flow and the average velocity drops by a factor of 2 (compare V_a from *ii* to *iii*). To confine and drive reagent, the belt is tensioned between a feed and collector roller (3 km long) that is corona-treated to achieve proper wetting (Fig. 1a inset right and Extended Data Fig. 2). The PET belt moves < 4 µm vertically (z-direction) at the maximum velocity (V_s = 0.32 m/s), and never touches or damages the biochip (cf. Extended Data Fig. 3). Overall, we create a thin and stable flow region with just two surfaces, as opposed to full solid wall confinement in flow cells.

Accurate base calling requires high reagent purity (< 0.1% cross-contamination of regeneration reagent) at concentrations >0.25 mM for dNTP. In replacement experiments, we can see the typical parabolic flow profile in the flow cell versus an approximately straight advection front in the r2r-fl cell (cf. Fig. 1c). Quantifying the time required to reach 99.9% sequencing reagent replacement, the flow cell takes ~ 40 s whereas for r2r-fl it is ~1 s (Fig. 1d), and with much lower required flushing volume (see top x-axis, and "Reagent replacement in the flow cell and r2r-fl" in the Method section to the end of this document). Overall, sequencing by r2r-fl improves performance and lowers reagent consumption (currently ~90% of the cost, see pie chart in Fig. 1e) from several times to about two orders of magnitude as compared to current flow-cell bases platforms. In fact, current DNA sequencing biochips are manufactured by the mature 28-nm-process at ~1000 US$ per 8" silicon wafer. Our attachment points are 220 nm in diameter and their pitch is 500 nm (Extended Data Fig. 4), resulting in a biochip cost of only US$ 0.07 per Gb (200 cycles). Based on the cost structure and reagent saving, we can achieve < US$ 15 per human genome using our r2r-fl approach (90 Gb). Currently, the second iteration of our r2r-fl sequencer is functioning in the China Genebank.

Comparison of roll-to-roll fluidics and flow cells by CFD simulation

We used computational fluid dynamics to confirm the drastic improvement in mass transfer speed and efficiency. In the r2r-fl cell (Fig. 2a), the reagent front stretches from a few mm at the entrance (Fig. 2b, left) to ~cm towards the end (Fig. 2b, right). In agreement with the replacement experiments (Fig. 1d), the time to fully replace the buffer with reagent is ~1 s. A 7 cm long single-lane model flow cell (see Fig. 2c for an open-case sequencer)—using shorter than normal inlet tubing of 30-cm, with identical height (20 µm) and average flow velocity (V_a = 0.16 m/s)—generates a pressure of 328 kPa (>>90 kpa) which causes leakage in practice and a deformation of 10 µm, which would lead to fracture of the top glass or UV glue (Fig. 2d and Extended Data Fig. 1a). The time to replace the buffer with reagent is ~30 s (Fig. 2e). See the comparison of simulation and experiments in Supplementary Video 1, and additional

CFD simulations for startup of r2r-fl, and reagent replacement at V_a = 0.015, 0.15, and 0.35 m/s in Extended Data Fig. 5 and 6.

When varying the average flow velocity, chip length or chamber height, we calculate that in r2r-fl the pressure drop Δp is always three orders of magnitude lower (Fig. 2f-h, respectively). Importantly, unlike in conventional flow cells, pressure variations in r2r-fl are invariant with their length, allowing flexible scaling of r2r-fl biochip (cf. Fig. 1a): the total area reaches 1.35 × 1.2 m (or >100 biochips of 7×7 cm^2 each) in a second generation r2r-fl setup for spatial omics (see Extended Data Fig. 7). The key advantage of r2r-fl is that the volume required to exchange 99.9% of the previous solution is only 60% more than the chamber volume at V_a = 0.16 m/s. In contrast, for the conventional flow cells it is 5 to 85 times greater than r2r-fl, depending on V_a and L (Fig. 2i, j). Particularly, for small flow cell systems like the Miseq (1 cm long biochip), the (inlet) tubing volume is more than an order of magnitude larger than the chamber. This requires large flushing volumes, unnecessarily increasing the dNTP consumption (Fig. 2k), and is thus responsible for the much higher cost per Gb. Overall, r2r-fl removes the constraints of traditional pressure-driven fluidics, since its plane Couette-flow has a negligible pressure drop over the fluidic cell, and therefore can be used from small (~ 1 cm^2) to large (> 1 m^2) biochip areas without modifications to the roll-to-roll machine. Moreover, the reagent consumption in r2r-fl is one order of magnitude lower, and reagent replacement can be achieved in under 2 s, thus enabling time-sensitive (e.g., emergency for a single patient) sequencing without compromising its cost-efficiency.

High-speed low-cost DNA sequencing by roll-to-roll fluidics

To validate r2r-fl for NGS applications, we sequence *E.coli*, SARS-CoV-2, and the human genome NA12878, as representatives for bacteria, viruses and animals, respectively (see full r2r-fl operation, biochemistry, and detection in Supplementary Video 2). Each of the five reagents is applied onto the PET belt by a dedicated slot die, according to the reaction sequence of SBS [30,31] (Fig. 3a and the Methods section). Once high-purity (>99.9%) reagents are sheared into the r2r-fl cell, we halt the belt movement

and increase the temperature to 57 ℃ for enzymatic reactions with DNA (Fig. 3b). To ensure sufficient fluorescence intensity, each single DNA fragment is amplified ~250 times to form a DNA nanoball (DNB) that is bright enough to be detected (Fig. 3c). [30,32] The nanowell size is set to ensure occupancy of only one DNB, stabilized via ionic bonding (Fig. 3d and e). To guarantee that all 250 sites in a DNB are terminated during SBS, we introduce an additional step in BGI chemistry to react with a highly reactive non-fluorescent synthesis reagent, COLD, after the fluorescence synthesis of HOT[13] (Fig. 3f). And then, the biochip is washed (WB2 in Fig. 3f) and transferred to the imaging system (Fig. 3g) for base calling and returned to the r2r-fl module for regeneration (RR and RB in Fig. 3f). Our largest single biochip of 15×15 cm^2 (Fig. 3h) generates up to 49,000 million reads. For a gap height $h = 20$ μm, reagent consumption is estimated to be 10,000 dNTPs per nanowell (Fig. 3i). We use a 2-second rinsing time to achieve 99.9% reagent purity (with a safety factor of 2 to ensure minimal cross-contamination) without experiencing a surge of run on. We achieved paired-end 100 bp (PE100) within a turnaround time 12 hours using a 7×7 cm^2 biochip. Both reagent consumption and washing speed were 30 times more efficient compared to the flow cell based sequencer with the same form factor.

Data quality comparison of NGS by roll-to-roll fluidics versus flow cell

We also compared the data quality of r2r-fl sequencing, confirming its minimal impact on SBS biochemistry. Similar to current flow cell sequencing technologies such as Illumina Novaseq X, using the human genome (NA12878) as a sample, the PE100 biochip production rate is 69.82% of the total area. Its heatmap demonstrates uniform and high sequencing data quality and high throughput, similar for *E.coli* and SARS-Cov-2 (Fig. 4a-c, Extended Data Figs. 8 & 9, Supplementary Tables 1-3). Fluorescent base intensity did not change significantly, and lag and run on were both within acceptable ranges (less than 20%) after 200 cycles (Fig. 4d-f). This indicates that DNBs are not damaged, biological reactions are nearly fully completed and contamination is minimal, ensuring accurate base calling. The four channels for A, T, C, and G all have good spectral separation, and the proportion of each base is consistent with existing commercial platforms (Fig. 4g). The sequencing accuracy, defined as the

probability of calling an incorrect base in 1000 bases (Q30), still requires further optimization of the bioreaction conditions but is currently always greater than 90% (Fig. 4h).

We conduct bioinformatics analysis to ensure the accuracy of the sequencing data. Compared with the human genome (hs37d5), [33] our sequencing results shows a mapping rate of 99.9%. Compared with the high-confidence data set of NA12878/HG001 in VCF format from GIAB (the Genome in a Bottle), [34] the precision of single nucleotide polymorphism (SNP) variants called from our sequencing results is 99.9%, and the sensitivity is 99.3%. The accuracy of insertions and deletions (INDEL) has a precision of 99.2% and a sensitivity of 98.0%, as presented in whole genome sequencing (WGS) data analysis (Fig. 4i, Supplementary Tables 4-8). In addition to WGS, we also demonstrate high data quality for *E. coli* sequencing (99.9% mapping rate) and high-precision sequencing results for SARS-CoV-2 with 0 nucleotide deletion, insertion or substitution compared to the literature (Extended Data Fig. 10, Supplementary Tables 9-12). [35] Thus, we conclude that the data quality of our r2r-fl platform meets or exceeds industry standards for sequencing. [36]

Discussion

Overall, roll-to-roll fluidics drastically advances next-generation sequencing (NGS), effectively addressing the hydromechanical constraints that have limited traditional pressure-driven flow cells. We have shown remarkable improvements in turnaround time, cost and throughput, enabling highly accurate sequencing for the human genome, *E. coli*, and SARS-CoV-2, with exceptional precision towards single nucleotide polymorphism (SNP). Our roll-to-roll methodology offers a faster and more affordable alternative for all aspects of gene work, allowing researchers to design comprehensive assays without making trade-offs between the breadth, depth, and frequency of genomic sequencing such as large-scale single-cell sequencing [37,38] or spatial omics [39,40] , clinicians to effectively respond to pathogens cancers [41,42] or genetic diseases [43,44], and policymakers to formulate timely public health strategies [45,46] . Beyond next generation sequencing, r2r-fl is a general fluid management concept for

fast, cost-efficient and scalable (micro)fluidic applications, with easier quality control in manufacturing, and holds immense promise in various healthcare domains and beyond.

Acknowledgements

This work is supported by BGI-Shenzhen and MGI Tech Co., Ltd.; the China Postdoctoral Science Foundation (2019M650215) awarded to Yanzhe Qin; China National Gene Bank; the National Key R&D Program of China (2020YFA0210800, led by Huanghao Yang, and 2020YFA0709900, led by Qiushui Chen); the National Natural Science Foundation of China (22027805, 22334004, and 22421002, led by Huanghao Yang, and 62134003 and U2441224, led by Qiushui Chen); and the Natural Science Foundation of Fujian Province (2022J06008, led by Qiushui Chen).

We thank Chutian Xing, Shengtao Lv for discussion and suggestion in engineering issues, Meihua Gong, Jing Yang and Xin Huang for instruction and guidance in biochemical process, Ming He, Mingyou Cao, Zhiheng Zhen and Yi Huang for discussion and calibration of imaging system, Haizhou Zhang for heating subsystem setting-up, Zihua Niu for liquid dynamic discussion, Wangsheng Li and Dandan Nie for biochip preparation and loading, Hong Xu for thickness determination, Mei Li for writing of Data Quality in the Method section, Jingjing Wang for discussion of model comparison, Ming Ni and Dong Wei for providing DNBSEQ-TX related resources, as well as Jian Liu, Hui Jiang, Xun Xu, Mengzhe Shen and Feng Mu for their funding support of this project. We thank Jeff Fredberg from Harvard T.H. Chan School of Public Health for edits.

Author Contributions Y.Q. initiated the project, conceived the idea and lead the development team. Q.C., T.M.H. and H.Y. help with paper writing, making figures and videos. S.Z. helped with imaging on DNBSEQ-T10. R.M. oversaw electrical engineering and coding of the project. Z.L. oversaw mechanical engineering of the project. X.Y and X.H helped with biochemical experiments. Z.Y and Y.Z helped with the data analysis. H.L. and T.M.H. helped with CFD, C.X. helped with DNB loading of the biochips, and X.Z helped with discussion of the paper. Y.Q., S.K., Q.C., and T.M.H. wrote the manuscript. People in BGI-Shenzhen have conducted the project, designed and tested two iterations of prototypes, with assistance of people in MGI. X.Y. and Y.G. were interns of BGI-Shenzhen. X.L., X.Q and X. Z. help with paper writing and potential application discussion. All authors discussed the results.

Competing Interests BGI is developing r2r-fl for NGS sequencers. The other authors declare that they have no competing interests.

Table 1 | Comparison of throughput, price, and reagent use of flow cell based NGS platforms (accounting for reagent recycling) and roll-to-roll fluidics

	NovaSeq X	NovaSeq 6000		NextSeq 2000	MiSeq	R2r-fl
Type	10B	S1	S4	P2	V2	S / M / L
Size	270 mm^2 ×8 lanes	431 mm^2	1292 mm^2×4 lanes	170 mm^2	18 mm^2	S:100 mm^2 M:4225 mm^2 L:21025 mm^2
†Reagent per cm^2 of biochip	32 μL	28 μL	8 μL	30 μL	250 μL	2 μL

†Reagent usage from the script folder of demo data in https://basespace.illumina.com/. Only synthesis reagents are calculated, recycled reagents' volume is deducted, and buffers' volume is not considered.

Figure 1 | Fluidic implementation of roll-to-roll fluidics (r2r-fl). a, Scheme of the r2r-fl setup. See main text for descriptions of i–iv (not to scale). V_s = 0.32 m/s. The reagent and washing solutions remain in the gap due to the Laplace pressure of $\gamma(\cos(\alpha)+\cos(\beta))/h$, where γ is the surface tension and α, β are the contact angles, see Extended Data Fig. 2. **b**, A photo of a typical r2r-fl setup holding a single biochip (up to 15×15 cm^2). **c**, Fluorescent solution (green) flushing out water in each lane of a DNBSEQ-G400 flow cell (upper panel) and r2r-fl cell (lower panel). Both are top views. The asterisk indicates the outlets where we measure the fluorescent intensity in **d** using actual sequencing reagents, showing fast (~1 s) and efficient fluid displacement (2 µL/cm^2) for r2r-fl, compared to ~40s and 65 µL/cm^2 for a single lane flow cell. Data are presented as mean values ± s.d. of the fluorescence changes in selected areas (marked with an asterisk in panel e) over at least 3 flushing samples for both the single-lane flow cell and the r2r-fl. **e,** Price per mega base sequenced from NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata, and with a star the predicted cost of the current work.

Figure 2 | Comparison of flow cells and r2r-fl. a, Picture of a 7×7 cm biochip mounted in the r2r-fl cell. **b**, Velocity lines (white arrows) and relative concentration (rainbow color) in the r2r-fl cell just at the cell entrance (left, up to 4 mm of biochip) and side-views at four successive times (right panel, full biochip length). V_a = 0.16 m/s; the pressure is ~100 Pa (cf. Extended Data Figure 6f). **c**, Picture of partially disassembled sequencer with DNBSEQ-G400 flow cell (see magnified inset). **d,** Pressure distribution in the DNBSEQ-G400 flow cell (rainbow colors); 170-µm-thick cover glass deformation (blue-to-red colors). Both by CFD, see Extended Data Fig. 1f-h for meshing details. **e**, Relative reagent concentration in the flow cell for five successive time (side views). Both cells in **(a-e)** have L = 7 cm, h = 20 µm, and V_a = 0.16 m/s. The width (w) for the r2r-fl cell is 7 cm, and 5.2 mm for the flow cell. **f-h**, Dependence of pressure drop average flow velocity (V_a), chamber length (L), and gap height (h), respectively. The pressure limit for the flow cell (e.g., microvalves or gasket) is 90 kPa (dashed line). **i, j**, Dependence of exchange ratio on average flow velocity **(i)** and chamber length **(j)**. The exchange ratio is defined as the supplied reagent volume required to achieve 99.9% purity at the biochip surface divided by the chamber volume. **k,** Dependence of the number of dNTP molecules supplied to each nanowell, on the tubing volume relative to the chamber volume for the flow cell (amplification N = 250 for DNA nanoballs). The dashed line shows the exchange ratio in r2r-fl, which does not require tubing. See also Supporting Video 1.

Figure 3 | Biochemical aspects of NGS using r2r-fl. a, Scheme showing how five sequential reagent/washing bands traverse the biochip, each reagent is deposited onto the moving PET belt by a dedicated slot die. Order is HOT, COLD, WB2, RR, RB. See Supporting Movie 2. **b,** Side view schematic of an operational unit for a single biochip, where the chip-holder increases to 57 °C when using the "HOT" sequencing reagent. **c,** Workflow for preparation of the DNA nanoballs (DNBs). Single-stranded DNA fragments are extracted from target DNA and amplified 250 times through rolling circle amplification. Each fragment forms a nanoball. **d,** Selective attachment of DNBs to the patterned biochip through ionic bonding. **e,** Schematic of the fluorescent imaging process for base calling. **f,** BGI's SBS biochemistry to sequence a DNB. The sequencing reagent (HOT) contains fluorescent dNTP (N stands for A, T, C or G) molecules and performs a self-terminated reaction to extend one complementary base. The sequencing reagent (COLD) reaction contains non-fluorescent dNTP to extend one self-

terminated complementary base with reaction points that failed to react with HOT. The biochip is then washed with WB2 and base calling is done by fluorescence imaging. Regeneration Reagent (RR) removes the fluorescent and azide groups to enable the next synthesis reaction cycle and is washed by Regeneration Buffer (RB). **g,** After flushing with washing buffer 2 (WB2), four-channel fluorescent imaging detects which type of nucleotide is added to each DNB, as shown in the merged image and the inset. **h,** Picture of the smallest 1×1 cm^2 and largest 15×15 cm^2 sequencing biochip (of the current work), and a single channel 6.3×0.75 cm^2 flow cell (DNBSEQ-200). **i**, Bar chart performance comparison of flow cell (DNBSEQ T7 and Novaseq X) versus r2r-fl for speed, reagent consumption, and data size. TAT is turnaround time for PE 100.

Figure 4 | Data quality comparison of next-generation sequencing (NGS) using r2r-fl versus flow cells. a-c, Typical heatmaps for the proportion of calling one incorrect base in 1000 bases (Q30) for r2r-fl, BGI DNBSEQ-T7, and Illumina NovaSeq X flow cells. Each dot represents a square field of view (FOV) for microscopy and is color-coded based on Q30, which is the benchmark for NGS. In (**a**), a selected region used for image registration is enclosed in a circle. In (**b**, **c**)**,** Q30 distribution indicates non-uniformity at edges due to uneven reagent distribution. White spots in (**b**) represent glue posts supporting the cover glass to prevent deformation. **d**, Fluorescence intensity of all four-color channels for r2r-fl, and two channels (green and blue) with different combinations to indicate the four nucleotides for Illumina NovaSeq X. **e,** Lag dependency on cycle number, representing the fraction of reaction points on DNBs falling behind the current cycle number. **f,** Run on dependency on cycle number, indicating the fraction of reaction points on DNBs exceeding the current cycle number. **g,** Base distribution for each sequencing cycle. **h,** Fraction of reads exceeding a base call accuracy of 99.9% (Q30 or higher). **i,** Precision and sensitivity for single nucleotide polymorphism (SNP) variants and insertions and deletions (INDEL) for r2r-fl and flow cell. Panels **d-i**, flow cell means the NovaSeq X. Illumina data is sourced from https://basespace.illumina.com.

Reference

1. Margulies, M. et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380 (2005).
2. Holt, R.A. & Jones, S.J. The new paradigm of flow cell sequencing. *Genome Res.* **18**, 839-846 (2008).
3. Uhlen, M. & Quake, S.R. Sequential sequencing by synthesis and the next-generation sequencing revolution. *Trends Biotechnol.* (2023).
4. Lu, H., Stratton, C.W. & Tang, Y.W. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *J. Med. Virol.* **92**, 401 (2020).
5. Kozlov, M. Ebola outbreak in Uganda: how worried are researchers. *Nature* **597** (2022).
6. Tegally, H. et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438-443 (2021).
7. Keita, A.K. et al. Resurgence of Ebola virus in 2021 in Guinea suggests a new paradigm for outbreaks. *Nature* **597**, 539-543 (2021).
8. Dudas, G. et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**, 309-315 (2017).
9. Fuller, C.W. et al. The challenges of sequencing by synthesis. *Nat. Biotechnol.* **27**, 1013-1023 (2009).
10. Balasubramanian, S. Sequencing nucleic acids: from chemistry to medicine. *Chem. Commun.* **47**, 7281-7286 (2011).
11. Rodriguez, R. & Krishnan, Y. The chemistry of next-generation sequencing. *Nat. Biotechnol.* (2023).
12. Case, D.J., Liu, Y., Kiss, I.Z., Angilella, J.-R. & Motter, A.E. Braess's paradox and programmable behaviour in microfluidic networks. *Nature* **574**, 647-652 (2019).
13. Bruus, H. Theoretical Microfluidics. (Oxford University Press, 2007).
14. Eisenstein, M. Innovative technologies crowd the short-read sequencing market. *Nature* **614**, 798-800 (2023).
15. Mohammed, M.I., Haswell, S. & Gibson, I. Lab-on-a-chip or Chip-in-a-lab: Challenges of Commercialization Lost in Translation. *Procedia Technology* **20**, 54-59 (2015).
16. Schwarze, K. et al. The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genet. Med.* **22**, 85-94 (2020).
17. Chai, J.H. et al. Cost-benefit analysis of introducing next-generation sequencing (metagenomic) pathogen testing in the setting of pyrexia of unknown origin. *PLoS One* **13**, e0194648 (2018).
18. Schermelleh, L. et al. Super-resolution microscopy demystified. *Nat. Cell Biol.* **21**, 72-84 (2019).
19. Heintzmann, R. & Huser, T. Super-resolution structured illumination microscopy. *Chem. Rev.* **117**, 13890-13908 (2017).
20. Ma, W. et al. Gene sequencing reaction device, gene sequencing system, and gene sequencing reaction method. *US11857973B2* (2024).
21. Beckett, N. & Caswell, N. Implementing barriers for controlled environments during sample processing and detection. *US20210354126A1* (2022).
22. Barbee, K. et al. Methods for biological sample processing and analysis. *US11499962B2* (2019).
23. Pennisi, E. $100 genome? New DNA sequencers could be a 'game changer' for biology, medicine. *Science*.
24. Reynolds, O. IV. On the theory of lubrication and its application to Mr. Beauchamp tower's experiments, including an experimental determination of the viscosity of olive oil. *Philos. Trans. R. Soc. London*, 157-234 (1886).
25. Reichardt, H. Über die Umströmung zylindrischer Körper in einer geradlinigen Couetteströmung. *(No Title)* (1954).

26. Tillmark, N. & Alfredsson, P.H. Experiments on transition in plane Couette flow. *J. Fluid Mech.* **235**, 89-102 (1992).
27. Krug, D., Lüthi, B., Seybold, H., Holzner, M. & Tsinober, A. 3D-PTV measurements in a plane Couette flow. *Exp. Fluids* **52**, 1349-1360 (2012).
28. Piau, J., Bremond, M., Couette, J. & Piau, M. Maurice Couette, one of the founders of rheology. *Rheol. Acta* **33**, 357-368 (1994).
29. Fardin, M., Perge, C. & Taberlet, N. "The hydrogen atom of fluid dynamics"–introduction to the Taylor–Couette flow for soft matter scientists. *Soft Matter* **10**, 3523-3535 (2014).
30. Mardis, E.R. DNA sequencing technologies: 2006–2016. *Nature protocols* **12**, 213-218 (2017).
31. Rodriguez, R. & Krishnan, Y. The chemistry of next-generation sequencing. *Nat. Biotechnol.* **41**, 1709-1715 (2023).
32. Ali, M.M. et al. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. *Chem. Soc. Rev.* **43**, 3324-3341 (2014).
33. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).
34. Krusche, P. et al. Best practices for benchmarking germline small-variant calls in human genomes. *Nat. Biotechnol.* **37**, 555-560 (2019).
35. Xiao, K. et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* **583**, 286-289 (2020).
36. Foox, J. et al. Performance assessment of DNA sequencing platforms in the ABRF Next-Generation Sequencing Study. *Nat. Biotechnol.* **39**, 1129-1140 (2021).
37. Zheng, W. et al. High-throughput, single-microbe genomics with strain resolution, applied to a human gut microbiome. *Science* **376**, eabm1483 (2022).
38. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865-868 (2017).
39. Lewis, S.M. et al. Spatial omics and multiplexed imaging to explore cancer biology. *Nat. Methods* **18**, 997-1012 (2021).
40. Bressan, D., Battistoni, G. & Hannon, G.J. The dawn of spatial omics. *Science* **381**, eabq4964 (2023).
41. Chiu, C.Y. & Miller, S.A. Clinical metagenomics. *Nature Reviews Genetics* **20**, 341-355 (2019).
42. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *New Engl. J. Med.* **382**, 727-733 (2020).
43. Church, A.J. et al. Molecular profiling identifies targeted therapy opportunities in pediatric solid cancer. *Nat. Med.* **28**, 1581-1589 (2022).
44. Nogrady, B. How cancer genomics is transforming diagnosis and treatment. *Nature* **579**, S10-S11 (2020).
45. Saunders, C.J. et al. Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. *Science translational medicine* **4**, 154ra135-154ra135 (2012).
46. Sazonovs, A. et al. Large-scale sequencing identifies multiple genes and rare variants associated with Crohn's disease susceptibility. *Nat. Genet.* **54**, 1275-1283 (2022).

Methods

Genomic materials. The genomic samples used in this study were obtained from the following sources: i) *E. col*i samples were derived from the MGI Standard Library Kit V3.0 (Part Number 1000005033); ii) human genome samples were obtained from the NIGMS Human Genetic Cell Repository (NA12878) by Coriell Institute, using the MGIEasy PCR-Free DNA Library Prep Set V1.1 (16 RXN, Part Numbers 1000013452); iii) SARS-CoV-2 genome samples were sourced from the SARS-CoV-2 Molecular Controls Kit - Full Genome (Part Numbers GW-CRBM002) by GeneWell Biotechnology Co., Ltd; and MGI ATOPlex RNA Multiplex PCR-based Library Preparation Set V3.0 (16 RXN, Part Numbers 940000132-00).

CFD simulations. COMSOL 6.0 was utilized for conducting fluid simulations in our study. We assumed standard Newtonian fluids with a viscosity of μ=10^{-3} Pa·s, and the diffusion coefficient of biomolecules was set to D=10^{-9} m^2/s. The mass transfer and biomolecule diffusion phenomena were calculated using the Navier-Strokes equation $\rho(\partial_t \boldsymbol{V} + (\boldsymbol{V}\cdot\nabla)\boldsymbol{v}) = -\nabla p + \mu\nabla^2\boldsymbol{V}$ and the convection-diffusion equation $\partial_t C + \boldsymbol{V}\cdot\nabla C = D\nabla^2 C$. To calculate the deformation of the cover glass, we applied the equations of motion $0 = \nabla\cdot(\mathrm{FS})^T + F_v$ and $F = \mathrm{I} + \nabla u_{solid}$. In a typical experiment, the biochip size was set to 7 cm × 7 cm for roll-to-roll fluidics (r2r-fl). The belt velocity was set to V_s = 0.32 m/s, resulting in an average flow velocity V_a = 0.16 m/s, and the gap height h = 20 μm. The model flow cell (see Extended Data Fig. 1f–h) was configured with a 30-cm inlet tubing, and the dimensions of the biochip were 72.3 mm in length and 5.2 mm in width.

We used standard quadrilateral meshing for our simulations. The flow cell chamber was configured with predefined normal settings calibrated for fluid dynamics. The inlet and outlet tubing had free quad bottoms, with 50 and 20 layers respectively, and element ratios of 100 and 20, symmetrically distributed. The biochip had 8 layers with a free quad bottom set to "finer" and calibrated for fluid dynamics. The outer surfaces of the tubing edges of the biochip had a boundary layer with 4 layers. For the r2r-fl, we used the following mesh size: the upper surface consisted of free quadrilateral nodes with a maximum size of 1 cm, swept for 6 layers with an element ratio of 6 and symmetric distribution.

Fabrication of biochips. Biochips comprising arrays of 2.5 × 10^{10} nanowells were fabricated from 7 cm × 7 cm silicon wafers utilizing deep ultraviolet lithography. In this process, hexamethyldisilane (HMDS) was vaporized and carefully deposited on the oxidized silicon substrate. HMDS formed a robust bond with OH groups on the silicon surface, resulting in the formation of a hydrophobic surface due to the attachment of methyl groups. These methyl groups remained on the surface, ensuring strong adhesion with the photoresist material. To achieve a desired thickness of 500 nm, the photoresist was spin-coated onto the wafers. Employing deep ultraviolet (DUV) lithography, a 15 × 15 cm^2 array of 10 × 10^{10} circular wells with a diameter of approximately 250 nm and a center-to-center separation of 500 nm was precisely created to enable one-to-one attachment of DNA nanoballs (DNBs). To eliminate any remnants of the photoresist from the bottom of the nanowells, the biochip underwent a thorough washing procedure involving acetone, followed by oxygen dry etching. Furthermore, through the application of the chemical vapor deposition (CVD) technique and amination of the silicon surface of the biochips by exposing them to HMDS vapor, the methyl groups on the surface were effectively converted into amine groups. This transformation rendered the nanowells hydrophilic and positively charged, facilitating the attachment of DNBs. Following the removal of all photoresist material, the biochip exhibited hydrophobic properties, except for the nanowells. Subsequently, a laser was employed to cut the wafer into the desired design shape, without necessitating any additional polishing steps.

Finally, optional 20 µm thick spacers, composed of 10 µm 3M UCT-10 double-stick tape and a 10 µm stainless steel shim, were affixed to the edges of the biochip. Handles, necessary for manipulation by a robotic arm, were attached using 3M adhesives.

Preparation of DNA nanoballs (DNBs). DNA nanoballs (DNBs) were prepared for various genomes[47], including the SARS-CoV-2 virus genome, human genome (after the aforementioned extraction steps), and *E. coli* from the Standard Library Kit V3.0 (Part Number 1000005033). The preparation was performed using the BGISEQ-500RS DNB Make Load Reagent (Part Number 1000005488) from MGI Tech Co. Ltd. The following steps were conducted: Initially, DNA was annealed using a Bio-Rad S1000TM thermocycler with temperature cycles of 95°C for 1 min, followed by 65 °C for 1 min, and 40 °C for 1 min [48]. Subsequently, 40 µL of Make DNB Enzyme Mix I and 4 µL of Make DNB Enzyme Mix II were added for rolling circle amplification, which took place at 30°C for 20 min. To terminate the amplification, the temperature was lowered to 0°C, and 20 µL of Make DNB Stop Buffer was added. This procedure resulted in the generation of 250 faithful copies of the libraries that formed DNA nanoballs (DNBs) with a normal size of ~ 220 nm (ref. [58]) and a concentration of 15-20 ng/µL.

Loading biochips with DNBs. To facilitate the adhesion of DNA nanoballs (DNBs) onto the wells of the biochip, the biochip was carefully placed into a loader, which functioned as a flow chamber with an 80 µm gap height set at 25 °C. The DNB solution was prepared by combining 400 µL of DNB Reagent, 500 µL of DNB Load Buffer I, and 300 µL of DNB Load Buffer II. To ensure uniform coating of the biochip's surface, a total of 1480 µL of the DNB solution was prepared, taking into account a reagent loss of 410 µL as bypass volume from the previous section into the loader. Denatured-free phi29 polymerase was adsorbed onto the aminated spots of the biochip, while Pluronic F68 (PF68), a PEO-PPO-PEO tri-block copolymer, prevented DNBs from connecting to the hexamethyldisilazane (HMDS) field through steric hindrance. PF68 contains hydrophilic hydroxyl groups at both ends (PEO part) and hydrophobic methyl groups in the middle (PPO part). When PF68 dissolved in DNB Load Buffer II flowed onto the biochip surface along with the DNA nanoballs, the principle of "like dissolves like" came into play. Due to electrical interactions between polar molecules, non-polar molecules in polar solvents tend to aggregate to minimize the hydrophobic area. Consequently, the hydrophobic PPO part of PF68 is attracted to the methyl groups of HMDS, causing the negatively charged DNBs to attach singly to the positively charged nanowells. To remove excess DNBs, a rinse with DNB Rinse Buffer (DRB) was performed. Subsequently, DNBs were condensed using DNB Crash Buffer (DCB). The biochips were then flushed with DNB Read Buffer (DRB) and Protein Washing Buffer (PWB). DRB facilitated the binding between DNBs and the biochip, while bovine serum albumin (BSA) in PWB was adsorbed by DNBs to provide support. To enhance stability, the biochips were flushed with DCB and REB once again. After these steps, the biochips were treated with Hybridization Buffer (HYB) and soaked in 1 µM AD153 Sequencing Primer 1 V3.0 at different temperatures (55°C for 9 minutes, 40°C for 2 minutes, and 25°C for 2 minutes). Based on the principle of complementary base pairing, the primers were spliced onto the corresponding single-stranded fragments on the DNB beads, enabling subsequent sequencing. To evaluate the loading effect, an initial round of sequencing was performed on the loader. After the sequencing primers hybridized to the adapter region of the DNBs, a fluorescently labeled dNTP probe was incorporated using cPAS DNA polymerase. Any unbound dNTP probes were then washed away. The DNB Flow Cell was imaged, and the fluorescence signal was converted to a digital signal. The base information was determined using MGI's proprietary base-calling software. To prevent the DNB-loaded biochips from drying out, they were stored in Wash Buffer 2 (WB2). All

reagents used in these processes were sourced from the DNBSEQ-T10x4 RS High-Throughput Sequencing Reagent Kit (Part Number 1000005488 from MGI Tech Co. Ltd).

Reagents for biochemistry. For sequencing, we conducted the standard Massively Parallel Sequencing (MPS) process using the DNBSEQ-T10x4 RS High-Throughput Sequencing Reagent Kit (category number 1000023411, MGI Tech Co. Ltd) designed for FCL PE100 sequencing. The process involves three distinct biochemical reactions as follows: 1) to replicate the complementary bases of the DNB template, we utilized Sequencing Reagent (HOT), which consists of cPAS DNA polymerase and dNTP molecules (N = A, T, C, or G). The dNTP molecules are conjugated with fluorophores that can be spectrally resolved by the imaging system[15]; 2) for bases of the DNB template that did not react with Sequencing Reagent (HOT), we employed non-fluorescent Sequencing Reagent (COLD), comprising unlabeled dNTPs and cPAS DNA polymerase for complementary base replication; and 3) To remove the fluorescent cap in preparation for the next sequencing cycle, we utilized Regeneration Reagent 2. The kit provides specific buffer washing solutions for each of these three reagents, including Wash Buffer, Sequencing Reagent, and Regeneration Buffer. It is important to note that unlike conventional NGS, there are no concerns regarding cross-contamination, allowing the elimination of the buffer for Sequencing Reagent (HOT) and direct flow of non-fluorescent Sequencing Reagent (COLD) immediately after Sequencing Reagent (HOT). To prevent beading up on the biochip and PET belt, the surface tension of all six solutions was reduced to less than 42 mN/m using Tween-20 surfactant. The first strand sequencing occurs from cycle 1 to cycle 100. Upon completion of the 1st strand sequencing, the 2nd strand generation primers and a polymerase with strand displacement activity are introduced to initiate 2nd strand synthesis. The polymerase extends the new primer until it reaches the original sequenced strand, displacing the original sequencing strand and forming a new single-stranded template. To maximize the length of the strand while ensuring its attachment to the original DNB, the newly generated 2nd strand is optimized. Following hybridization of the 2nd strand sequencing primer through multiple displacement amplification (MDA), the same sequencing chemistry employed during 1st strand sequencing is used for 2nd strand sequencing. The newly generated 2nd strand template contains a higher number of copies of the insert DNA, resulting in stronger signals and increased sequencing accuracy.

Roll-to-roll fluidics (r2r-fl) platform setup. To establish r2r-fl, a custom slot-die coater and the biochip holder from DNBSEQ-T7 were integrated onto an industrial roller (Siemens 6SL3210-5FE10-4UA0). The setup included five slot-die coating units (Foshan Edge Development Mold Mechanical Technology, Inc) corresponding to the five reagent formulations discussed previously. The coating roller served as the driving mechanism, while the tension of the PET belt was carefully adjusted to 20 kg to ensure flatness. The belt speed was set to 0.32 m/s, and the slot-die flow rate was meticulously adjusted to achieve a 20-μm gap thickness. To render the belt hydrophilic, a corona pretreatment was employed. The PET belt had a thickness of 100 μm and a total length of 3000 m, allowing for up to 500 sequencing cycles.

Roll-to-roll fluidics platform operation. A Robotic Arm, C8XL by Epson, facilitated the movement of the biochip between the r2r-fl platform and the imaging subsystems, which was adapted from the commercial DNBSEQ-T10 by BGI. During the intervals between processing cycles, the biochip was securely stored in a dedicated transfer slot containing Wash Buffer 2. The biochemical processes involved in base calling comprised two steps for base addition and one step for fluorescent group cleavage, necessitating three reagent bands and two washing bands. Prior to operation, the slot dies were primed to ensure optimal performance. Each band had a length of 40 cm, which was more than twice the size of the biochip, ensuring complete replacement of the reagents. A complete biochemical cycle

was accomplished within 60 seconds. The transfer of the biochip from the transfer slot bath to the imaging subsystem and back, facilitated by the robotic arm, was completed in 20 seconds. The imaging process itself required 60 seconds.

Reagent replacement in the flow cell and r2r-fl. The cross-contamination of dNTP synthesis reagents (HOT and COLD in BGI's SBS chemistry) and the regeneration reagent, which removed the steric hindrance group, should be avoided as much as possible. Otherwise, more than one base would be added to each reaction site and drive run on surge. In quality control standard, it is less than 0.1% mixing. In practice, we utilized two reagents (one of them was a buffer) to achieve this goal. To evaluate the cleaning efficiency for the flow cell and r2r-fl cell, we employed the standard method in R&D to compare the fluorescent replacement speed of the actual sequencing reagent (HOT). First, we filled a new flow cell or a r2r-fl cell with Wash Buffer 2, took pictures near the outlet using the imaging module of the sequencer, and considered the average gray value before any fluorescent reagent contamination as the background value. Next, we diluted HOT about twice until the average gray value was close to but less than the saturation value of 65535, considering it as the full fluorescent value. Then, we used Wash Buffer 2 to clean the HOT until the remaining gray value was less than "the background value + 0.1% x (full fluorescent value - background value)". For example, if the full fluorescent value of the flow cell is 63000, and its background value is 300, then its endpoint was considered the moment of gray value for the outlet reaching "300 + (63000-300) = 362.7". We used DNBSEQ-G50 and its flow cell with a 7 cm length and 50 μm gap height, along with a 7 × 7 cm^2 r2r-fl cell with 50 μm gap height for comparison. The flow cell was connected to the Wash buffer 2 bottle with a 30 cm tubing with a 0.8 mm inner diameter, shorter than the actual product of DNBSEQ-G50. For the gray value at a specific moment (e.g., 20 s) in the figure, we run the pump/PET belt for a certain time and stopped, took a picture using the imaging module, and recorded the average gray value close to the outlet on the biochip. Each data point was repeated at least three times.

Reagent layer uniformity and thickness. We achieved uniform coatings by carefully adjusting the surface energy of both the reagents and the polymeric belt. To promote spreading and prevent beading up, we employed corona pretreatment to increase the PET belt (Nanya plastic corporation, Rynite 935 BK505 with 35% mica/glass reinforced modified polyethylene terephthalate resin) surface energy. Additionally, we incorporated 0.5% Tween-20 to reduce the surface tension of the reagents. The thickness of the layers was measured using the Keyence sensor SI-T80, which offers a linearity of ±0.05% of full scale and a repeatability of 0.5 μm. To determine the actual liquid thickness, we moved the laser detector across the wet zone of the PET belt from one end to the other, resulting in a measured thickness of 10±2 μm with a variation of ± 0.5 μm.

Identification of COVID-19 variants. We conducted a comprehensive analysis to ensure the traceability of COVID-19 data on our platform, including the identification of variants. The COVID-19 sequence data obtained from our platform was aligned to the SARS-CoV-2 wild-type reference genome (Genbank: MN908947) using the BWA-MEM alignment tool. Variants in the alignment results were called using freebayes, and the results were further annotated using SnpEff for comprehensive variation calling. To validate the accuracy of the called variants, we compared them with the reference genome and generated a FASTA file containing the sequenced library with the identified variants. Notably, our analysis revealed no variants, indicating that the sequence of the library perfectly matched the reference genome. To further confirm the traceability, we utilized Nextclade [49] (https://clades.nextstrain.org) to track the FASTA file of the sequenced library. The traceable results demonstrated that the sequence

library exhibited a high degree of homology with the SARS-CoV-2 wild-type reference genome (Genbank: MN908947).

Whole genome sequencing (WGS) analysis. The data were processed using SOAPnuke [50] and aligned to the human reference genome (hs37d5) using the BWA-MEM [51] alignment algorithm. SAMtools was then utilized to sort the alignment results [52], and Picard was employed to mark duplicates after the sorting process [53]. GATK3.7 was applied to correct the base quality values using the BQSR (Base Quality Score Recalibration) algorithm, following the parameter recommendations provided by the official GATK website [54]. Germline mutations were called using GATK HaplotypeCaller. Copy number variations were analyzed using CNVnator [55], while structural variants were assessed using BreakDancer (V1.0) [56]. To provide comprehensive annotations, the identified variants were annotated using ANNOVAR and RTG vcfeval. These rigorous data processing and analysis steps utilizing established tools and algorithms ensure the accuracy and reliability of the obtained results.

Data quality. Lag and run-on are attributed to imperfections in the biochemical synthesis procedure. Lag indicates that one or more bases have failed to be synthesized during the current sequencing cycle, while run-on signifies premature synthesis of one or more bases. To identify instances of lag and run-on, we utilized BGI's commercial software, which also generates informative figures and provides various indicators such as Q30. In order to determine the success of the loading process, we evaluated the Q30 report.

Data availability. The Illumina sequencing quality data used in this work for comparison are publicly available through basespace (basespace.illumina.com, registration needed), sequencing data and analysis are available in "Yanzhe Qin. Fast, cost-effective, and flexible DNA sequencing by roll-to-roll fluidics [DS/OL]. V1. Science Data Bank, 1[2024-12-24]. https://doi.org/10.57760/sciencedb.18851. DOI:10.57760/sciencedb.18851." Other data that support the findings of this study are available from the corresponding author upon reasonable request.

Method References

47. Drmanac, R. et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* **327**, 78-81 (2010).
48. Li, Z. et al. DNB-based on-chip motif finding: A high-throughput method to profile different types of protein-DNA interactions. *Science advances* **6**, eabb3350 (2020).
49. Aksamentov, I., Roemer, C., Hodcroft, E.B. & Neher, R.A. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *Journal of Open Source Software* **6**, 3773 (2021).
50. Chen, Y. et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* **7**, gix120 (2018).
51. Jung, Y. & Han, D. BWA-MEME: BWA-MEM emulated with a machine learning approach. *bioRxiv* (2021).
52. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
53. Toolkit, P. (2019).
54. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*, 201178 (2017).
55. Abyzov, A., Urban, A.E., Snyder, M. & Gerstein, M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974-984 (2011).
56. Chen, K. et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* **6**, 677-681 (2009).

624

Fast, cost-effective, and flexible DNA sequencing by roll-to-roll fluidics

Yanzhe Qin[1,2,3*], Stephan A. Koehler[4], Yunyan Ling[1], Shiqiang Yu[1], Yongjie Zhang[3], Jie Luo[1], Kaijian Chen[1], Junjie Luo[1], Haiyang Chu[1], Junjie Zeng[1], Fei Wang[1], Wei Li[1], Dan Li[1], Xinran Yu[5], Xiangchao Wu[1], Shengming Zhao[2], Hao Lu[2], Ziqing Deng[2], Zhijian Yang[3], Ruibin Mai[2], Zhuo Liu[2], Zihua Niu[2], Xin Huang[1], Chengmei Xing[1], Wenwei Zhang[1], Xun Xu[1], Xingcai Zhang[6], Yongyou Hu[7], Xinwen Peng[7], Yiming Gong[8], Qiushui Chen[3*] , Xiaoxing Liao[9], Xiangmeng Qu[10], Thomas M. Hermans[11*]& Huanghao Yang[3*]

[1]*BGI-Shenzhen, Shenzhen 518083, China.*

[2]*MGI, BGI-Shenzhen, Shenzhen 518083, China.*

[3]*New Cornerstone Science Laboratory, College of Chemistry, Fuzhou University, Fuzhou 350002, China.*

[4]*Artificial Intelligence and Machine Learning Laboratory, Riverside Research Institute, 70 Westview St Suite 3, Lexington, MA 02421, USA.*

[5]*Harvard School of Public Health, Boston, Massachusetts 02115, United States.*

[6]*Department of Materials Science and Engineering, Stanford University, Stanford, CA, 94305, USA*

[7]*School of Light Industry and Engineering, School of Environment and Energy, South China University of Technology, Guangzhou 510006, China.*

[8]*Department of Physics, Kyoto University, Kyoto 606-8502, Japan.*

[9]*Emergency and Disaster Medical Center,The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen 518107, China*

[10]*School of Biomedical Engineering, Sun Yat-Sen University, Shenzhen 518107, China*

[11]*IMDEA Nanoscience, C/ Faraday 9, Madrid 28049, Spain.*

[*]Corresponding author. E-mail: yanzheqin@126.com (Y.Q.); thomas.hermans@imdea.org (T.M.H.); hhyang@fzu.edu.cn (H.Y.); qchen@fzu.edu.cn (Q.C.)

Despite decreasing prices per DNA base pair over the past decade, small-scale next-generation sequencing (NGS) remains prohibitive in urgent clinical settings, since the reagent efficiency of sequence-by-synthesis (SBS) is low in the corresponding flow cells. Here we show NGS by roll-to-roll fluidics (r2r-fl), which is cost-effective for flexible biochip sizes. R2r-fl is a practical implementation of plane-Couette flow, with up to 85 times lower reagent consumption (US$ 0.16 per giga base pair), rinsing times of less than 2 seconds, and shortened turnaround times (TAT) for paired-end 100 bp (PE100) sequencing from days to under 12 hours. We maintain over 99.9% precision and 99.3% sensitivity of single nucleotide polymorphisms (SNP) for the human genome, 99.9% mapping rate for *E.coli*, and minimal nucleotide substitutions, deletions, or insertions in

the SARS-CoV-2 Alpha strain. Overall, r2r-fl makes NGS drastically faster and more affordable, thus contributing to revolutions in dealing with pathogens, cancers, genetic diseases, and personalized medicine.

Introduction

High-throughput DNA sequencing[1-3] is increasingly common in public health management, and has proven to be an effective tool for responding to epidemics, such as providing early warnings of regional outbreaks for SARS-CoV-2 or Ebola, [4, 5] as well as tracking of their evolution. [6-8] The core component of the current next-generation sequencing (NGS) system is the flow cell that was pioneered in 2005. [1] Typical cells have a 50-100 µm gap height with a glass slide covering a patterned biochip onto which single-stranded DNA molecules are adhered (see Extended Data Fig. 1a for commercially available cells).[3,9] To perform cycling SBS—involving DNA extension by a dNTP (deoxy-ribonucleoside triphosphate), fluorescence detection ('base calling') and regeneration[10,11] (see the Methods section), repetitively—the reagents need to be fully (99.9%) replaced to ensure extending by only a single complementary base in each cycle. To achieve that in a flow cell system, one needs to flush the cell chamber and tubing with a volume equal to several times their internal volume, due to the parabolic Poiseuille flow profile (Extended Data Fig. 1b-e). The pressure drop over the chamber $\Delta p \approx \frac{12\mu L V_a}{h^2}$, depends on the chamber length L [m], height h [m], dynamic viscosity μ [N·s/m] and the average fluid velocity V_a [m/s]. [12, 13] The maximum operating pressure for current sequencers is about ~90 kPa (~0.9 bar) due to pressure members such as manifolds and O-rings, which limits further performance improvement, such as: i) throughput, since this scales with chamber length L; ii) reagent efficiency, since for smaller gaps h less reagent is wasted; iii) turnaround time (TAT), since greater flow velocity V_a allows for faster reagent cycling.

Currently, the highest throughput flow cell in NGS platforms can generate 8 Tb at an average price of US$ 2 per giga base pair ($/Gb), with a TAT of two days. [14] However, for smaller biochips the reagent efficiency is getting lower, and prices surge. The high operational costs of NGS are due to fabrication

of flow cells with low tolerance, [15] and especially the use of expensive reagents (>90% of the total price, same for manufacturers). [16, 17] To lower reagent usage, Illumina launched the Nextseq 2000 and NovaSeq X, which recycles 30% of the reagents and triples the density of DNA molecule attachment points on the biochip, resulting in improved data throughput and enhanced reagent efficiency. Unfortunately, its potential is limited by the minimum size for the DNA attachment points to fulfill requisite brightness (~100 nm), and requiring super-resolution imaging as attachment points are closer than the Rayleigh limit which significantly prolongs the biochip scanning process. [18, 19] An alternative approach to reduce cost is to expand the biochip area, such as DNBSEQ-T7 and NovaSeq 6000. However, this approach may extend patient waiting time, as economies of scale are only possible when enough samples are collected and loaded. To reduce TAT, small flow cells such as Miseq are available that have a throughput of 4.5 Gb for PE150. However, the reagent usage efficiency is compromised due to the significant flushing requirement of the supply tubing, whose volume is much larger than the flow cell chamber itself (Table 1). Overall, flow cell platforms are subject to hydromechanical constraints imposed by Poiseuille flow, and engineers have to make a trade-off between cost, throughput and TAT. To overcome these constraints, BGI[20] and Ultima[21, 22] genomics have borrowed technologies from the semiconductor industry such as RCA cleaning or spin-coating in order to reduce the reagent thickness on the wafer. These arrangements get rid of the flow cell structure, using a naked wafer instead, and both companies have announced $100 human genome sequencing[23]. However, to be cost-effective both systems need to collect enough (patient) samples to process large wafer(s) before starting a single run, with a data production which far exceeds daily clinical demand (see Supporting Information section ‘Discussion on flow-cell free DNA sequencers’). Currently, there is a pressing clinical need for sequencers that can economically process small numbers of samples in a single run, which today’s technology cannot adequately meet.

Plane-Couette flow (pCF)—i.e., negligible pressure-drop shear-flow between two infinite large moving parallel plates—was first put forward by Reynolds in 1884, in describing lubrication theory. [24]

Experimental implementations, mainly to study high Reynolds number (*Re*) turbulent flow transitions, had to wait until 1954 by Reichardt [25]. In the few experimental records on pCF, a looped belt tensioned between rollers is fully immersed in fluid, [26,27] rendering it impractical for implementations other than fundamental fluid dynamics study. In contrast, circular Couette flow—a shear flow between two concentric cylinders—was rapidly adopted in the following years by other scientific communities, e.g., as a basis for rheology. [28,29] However, next we will show a practical non-immersed implementation of pCF we call roll-to-roll fluidics (r2r-fl, Fig. 1), which has very efficient mass transfer irrespective of biochip size, and is ideal for rapid and flexible scale NGS which meet clinical demand with low cost, high speed and flexible sample sizes.

Results

Establishment of Plane-Couette flow by roll-to-roll fluidics

Specifically, we combine slot-die injection to layer reagents at the bottom side of a moving PET (Polyethylene terephthalate) belt that grazes a stationary biochip (Fig. 1a). The flow behavior is schematically explained in the inset, where: i) upon slot die injection, ii) a plug flow layer of height $h/2$ is formed, which iii) expands and wets the biochip, forming plane-Couette flow of height h at average velocity V_a, and subsequently iv) is dragged by the PET belt after the biochip, reestablishing plug flow. Thus, the plug flow layer (Fig. 1a, *ii*) only needs to be half the height of that in the pCF zone (*iii*), due to conservation of flow and the average velocity drops by a factor of 2 (compare V_a from *ii* to *iii*). To confine and drive reagent, the belt is tensioned between a feed and collector roller (3 km long) that is corona-treated to achieve proper wetting (Fig. 1a inset right and Extended Data Fig. 2). The PET belt moves < 4 µm vertically (z-direction) at the maximum velocity (V_s = 0.32 m/s), and never touches or damages the biochip (cf. Extended Data Fig. 3). Overall, we create a thin and stable flow region with just two surfaces, as opposed to full solid wall confinement in flow cells.

Accurate base calling requires high reagent purity (< 0.1% cross-contamination of regeneration reagent) at concentrations >0.25 mM for dNTP. In replacement experiments, we can see the typical parabolic flow profile in the flow cell versus an approximately straight advection front in the r2r-fl cell (cf. Fig. 1c). Quantifying the time required to reach 99.9% sequencing reagent replacement, the flow cell takes ~ 40 s whereas for r2r-fl it is ~1 s (Fig. 1d), and with much lower required flushing volume (see top x-axis, and "Reagent replacement in the flow cell and r2r-fl" in the Method section to the end of this document). Overall, sequencing by r2r-fl improves performance and lowers reagent consumption (currently ~90% of the cost, see pie chart in Fig. 1e) from several times to about two orders of magnitude as compared to current flow-cell bases platforms. In fact, current DNA sequencing biochips are manufactured by the mature 28-nm-process at ~1000 US$ per 8" silicon wafer. Our attachment points are 220 nm in diameter and their pitch is 500 nm (Extended Data Fig. 4), resulting in a biochip cost of only US$ 0.07 per Gb (200 cycles). Based on the cost structure and reagent saving, we can achieve < US$ 15 per human genome using our r2r-fl approach (90 Gb). Currently, the second iteration of our r2r-fl sequencer is functioning in the China Genebank.

Comparison of roll-to-roll fluidics and flow cells by CFD simulation

We used computational fluid dynamics to confirm the drastic improvement in mass transfer speed and efficiency. In the r2r-fl cell (Fig. 2a), the reagent front stretches from a few mm at the entrance (Fig. 2b, left) to ~cm towards the end (Fig. 2b, right). In agreement with the replacement experiments (Fig. 1d), the time to fully replace the buffer with reagent is ~1 s. A 7 cm long single-lane model flow cell (see Fig. 2c for an open-case sequencer)—using shorter than normal inlet tubing of 30-cm, with identical height (20 µm) and average flow velocity (V_a = 0.16 m/s)—generates a pressure of 328 kPa (>>90 kpa) which causes leakage in practice and a deformation of 10 µm, which would lead to fracture of the top glass or UV glue (Fig. 2d and Extended Data Fig. 1a). The time to replace the buffer with reagent is ~30 s (Fig. 2e). See the comparison of simulation and experiments in Supplementary Video 1, and additional

CFD simulations for startup of r2r-fl, and reagent replacement at V_a = 0.015, 0.15, and 0.35 m/s in Extended Data Fig. 5 and 6.

When varying the average flow velocity, chip length or chamber height, we calculate that in r2r-fl the pressure drop Δp is always three orders of magnitude lower (Fig. 2f-h, respectively). Importantly, unlike in conventional flow cells, pressure variations in r2r-fl are invariant with their length, allowing flexible scaling of r2r-fl biochip (cf. Fig. 1a): the total area reaches 1.35 × 1.2 m (or >100 biochips of 7×7 cm^2 each) in a second generation r2r-fl setup for spatial omics (see Extended Data Fig. 7). The key advantage of r2r-fl is that the volume required to exchange 99.9% of the previous solution is only 60% more than the chamber volume at V_a = 0.16 m/s. In contrast, for the conventional flow cells it is 5 to 85 times greater than r2r-fl, depending on V_a and L (Fig. 2i, j). Particularly, for small flow cell systems like the Miseq (1 cm long biochip), the (inlet) tubing volume is more than an order of magnitude larger than the chamber. This requires large flushing volumes, unnecessarily increasing the dNTP consumption (Fig. 2k), and is thus responsible for the much higher cost per Gb. Overall, r2r-fl removes the constraints of traditional pressure-driven fluidics, since its plane Couette-flow has a negligible pressure drop over the fluidic cell, and therefore can be used from small (~ 1 cm^2) to large (> 1 m^2) biochip areas without modifications to the roll-to-roll machine. Moreover, the reagent consumption in r2r-fl is one order of magnitude lower, and reagent replacement can be achieved in under 2 s, thus enabling time-sensitive (e.g., emergency for a single patient) sequencing without compromising its cost-efficiency.

High-speed low-cost DNA sequencing by roll-to-roll fluidics

To validate r2r-fl for NGS applications, we sequence *E.coli*, SARS-CoV-2, and the human genome NA12878, as representatives for bacteria, viruses and animals, respectively (see full r2r-fl operation, biochemistry, and detection in Supplementary Video 2). Each of the five reagents is applied onto the PET belt by a dedicated slot die, according to the reaction sequence of SBS [30,31] (Fig. 3a and the Methods section). Once high-purity (>99.9%) reagents are sheared into the r2r-fl cell, we halt the belt movement

and increase the temperature to 57 ℃ for enzymatic reactions with DNA (Fig. 3b). To ensure sufficient fluorescence intensity, each single DNA fragment is amplified ~250 times to form a DNA nanoball (DNB) that is bright enough to be detected (Fig. 3c). [30,32] The nanowell size is set to ensure occupancy of only one DNB, stabilized via ionic bonding (Fig. 3d and e). To guarantee that all 250 sites in a DNB are terminated during SBS, we introduce an additional step in BGI chemistry to react with a highly reactive non-fluorescent synthesis reagent, COLD, after the fluorescence synthesis of HOT[13] (Fig. 3f). And then, the biochip is washed (WB2 in Fig. 3f) and transferred to the imaging system (Fig. 3g) for base calling and returned to the r2r-fl module for regeneration (RR and RB in Fig. 3f). Our largest single biochip of 15×15 cm^2 (Fig. 3h) generates up to 49,000 million reads. For a gap height $h = 20$ μm, reagent consumption is estimated to be 10,000 dNTPs per nanowell (Fig. 3i). We use a 2-second rinsing time to achieve 99.9% reagent purity (with a safety factor of 2 to ensure minimal cross-contamination) without experiencing a surge of run on. We achieved paired-end 100 bp (PE100) within a turnaround time 12 hours using a 7×7 cm^2 biochip. Both reagent consumption and washing speed were 30 times more efficient compared to the flow cell based sequencer with the same form factor.

Data quality comparison of NGS by roll-to-roll fluidics versus flow cell

We also compared the data quality of r2r-fl sequencing, confirming its minimal impact on SBS biochemistry. Similar to current flow cell sequencing technologies such as Illumina Novaseq X, using the human genome (NA12878) as a sample, the PE100 biochip production rate is 69.82% of the total area. Its heatmap demonstrates uniform and high sequencing data quality and high throughput, similar for *E.coli* and SARS-Cov-2 (Fig. 4a-c, Extended Data Figs. 8 & 9, Supplementary Tables 1-3). Fluorescent base intensity did not change significantly, and lag and run on were both within acceptable ranges (less than 20%) after 200 cycles (Fig. 4d-f). This indicates that DNBs are not damaged, biological reactions are nearly fully completed and contamination is minimal, ensuring accurate base calling. The four channels for A, T, C, and G all have good spectral separation, and the proportion of each base is consistent with existing commercial platforms (Fig. 4g). The sequencing accuracy, defined as the

probability of calling an incorrect base in 1000 bases (Q30), still requires further optimization of the bioreaction conditions but is currently always greater than 90% (Fig. 4h).

We conduct bioinformatics analysis to ensure the accuracy of the sequencing data. Compared with the human genome (hs37d5), [33] our sequencing results shows a mapping rate of 99.9%. Compared with the high-confidence data set of NA12878/HG001 in VCF format from GIAB (the Genome in a Bottle), [34] the precision of single nucleotide polymorphism (SNP) variants called from our sequencing results is 99.9%, and the sensitivity is 99.3%. The accuracy of insertions and deletions (INDEL) has a precision of 99.2% and a sensitivity of 98.0%, as presented in whole genome sequencing (WGS) data analysis (Fig. 4i, Supplementary Tables 4-8). In addition to WGS, we also demonstrate high data quality for *E. coli* sequencing (99.9% mapping rate) and high-precision sequencing results for SARS-CoV-2 with 0 nucleotide deletion, insertion or substitution compared to the literature (Extended Data Fig. 10, Supplementary Tables 9-12). [35] Thus, we conclude that the data quality of our r2r-fl platform meets or exceeds industry standards for sequencing. [36]

Discussion

Overall, roll-to-roll fluidics drastically advances next-generation sequencing (NGS), effectively addressing the hydromechanical constraints that have limited traditional pressure-driven flow cells. We have shown remarkable improvements in turnaround time, cost and throughput, enabling highly accurate sequencing for the human genome, *E. coli*, and SARS-CoV-2, with exceptional precision towards single nucleotide polymorphism (SNP). Our roll-to-roll methodology offers a faster and more affordable alternative for all aspects of gene work, allowing researchers to design comprehensive assays without making trade-offs between the breadth, depth, and frequency of genomic sequencing such as large-scale single-cell sequencing [37,38] or spatial omics [39,40] , clinicians to effectively respond to pathogens cancers [41,42] or genetic diseases [43,44], and policymakers to formulate timely public health strategies [45,46] . Beyond next generation sequencing, r2r-fl is a general fluid management concept for

fast, cost-efficient and scalable (micro)fluidic applications, with easier quality control in manufacturing, and holds immense promise in various healthcare domains and beyond.

Acknowledgements

This work is supported by BGI-Shenzhen and MGI Tech Co., Ltd.; the China Postdoctoral Science Foundation (2019M650215) awarded to Yanzhe Qin; China National Gene Bank; the National Key R&D Program of China (2020YFA0210800, led by Huanghao Yang, and 2020YFA0709900, led by Qiushui Chen); the National Natural Science Foundation of China (22027805, 22334004, and 22421002, led by Huanghao Yang, and 62134003 and U2441224, led by Qiushui Chen); and the Natural Science Foundation of Fujian Province (2022J06008, led by Qiushui Chen).

We thank Chutian Xing, Shengtao Lv for discussion and suggestion in engineering issues, Meihua Gong, Jing Yang and Xin Huang for instruction and guidance in biochemical process, Ming He, Mingyou Cao, Zhiheng Zhen and Yi Huang for discussion and calibration of imaging system, Haizhou Zhang for heating subsystem setting-up, Zihua Niu for liquid dynamic discussion, Wangsheng Li and Dandan Nie for biochip preparation and loading, Hong Xu for thickness determination, Mei Li for writing of Data Quality in the Method section, Jingjing Wang for discussion of model comparison, Ming Ni and Dong Wei for providing DNBSEQ-TX related resources, as well as Jian Liu, Hui Jiang, Xun Xu, Mengzhe Shen and Feng Mu for their funding support of this project. We thank Jeff Fredberg from Harvard T.H. Chan School of Public Health for edits.

Author Contributions Y.Q. initiated the project, conceived the idea and lead the development team. Q.C., T.M.H and H.Y. help with writing paper, making figures and videos. S.Z. helped with imaging on DNBSEQ-T10. R.M. oversaw electrical engineering and coding of the project. Z.L. oversaw mechanical engineering of the project. X.Y and X.H helped with biochemical experiments. Z.Y and Y.Z helped with the data analysis. H.L. and T.M.H. helped with CFD, C.X. helped with DNB loading of the biochips, and X.Z helped with discussion of the paper. Y.Q., S.K., Q.C., and T.M.H. wrote the manuscript. People in BGI-Shenzhen have conducted the project, designed and tested two iterations of prototypes, with assistance of people in MGI. X.Y. and Y.G. were interns of BGI-Shenzhen. X.L., X.Q and X. Z. help with paper writing and potential application discussion. All authors discussed the results.

Competing Interests BGI is developing r2r-fl for NGS sequencers. The other authors declare that they have no competing interests.

Table 1 | Comparison of throughput, price, and reagent use of flow cell based NGS platforms (accounting for reagent recycling) and roll-to-roll fluidics

	NovaSeq X	NovaSeq 6000		NextSeq 2000	MiSeq	R2r-fl
Type	10B	S1	S4	P2	V2	S / M / L
Size	270 mm^2 ×8 lanes	431 mm^2	1292 mm^2×4 lanes	170 mm^2	18 mm^2	S:100 mm^2 M:4225 mm^2 L:21025 mm^2
†Reagent per cm^2 of biochip	32 μL	28 μL	8 μL	30 μL	250 μL	2 μL

†Reagent usage from the script folder of demo data in https://basespace.illumina.com/. Only synthesis reagents are calculated, recycled reagents' volume is deducted, and buffers' volume is not considered.

Figure 1 | Fluidic implementation of roll-to-roll fluidics (r2r-fl). a, Scheme of the r2r-fl setup. See main text for descriptions of i–iv (not to scale). V_s = 0.32 m/s. The reagent and washing solutions remain in the gap due to the Laplace pressure of $\gamma(\cos(\alpha)+\cos(\beta))/h$, where γ is the surface tension and α, β are the contact angles, see Extended Data Fig. 2. **b**, A photo of a typical r2r-fl setup holding a single biochip (up to 15×15 cm^2). **c**, Fluorescent solution (green) flushing out water in each lane of a DNBSEQ-G400 flow cell (upper panel) and r2r-fl cell (lower panel). Both are top views. The asterisk indicates the outlets where we measure the fluorescent intensity in **d** using actual sequencing reagents, showing fast (~1 s) and efficient fluid displacement (2 µL/cm^2) for r2r-fl, compared to ~40s and 65 µL/cm^2 for a single lane flow cell. Data are presented as mean values ± s.d. of the fluorescence changes in selected areas (marked with an asterisk in panel e) over at least 3 flushing samples for both the single-lane flow cell and the r2r-fl. **e,** Price per mega base sequenced from NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata, and with a star the predicted cost of the current work.

Figure 2 | Comparison of flow cells and r2r-fl. a, Picture of a 7×7 cm biochip mounted in the r2r-fl cell. **b**, Velocity lines (white arrows) and relative concentration (rainbow color) in the r2r-fl cell just at the cell entrance (left, up to 4 mm of biochip) and side-views at four successive times (right panel, full biochip length). V_a = 0.16 m/s; the pressure is ~100 Pa (cf. Extended Data Figure 6f). **c**, Picture of partially disassembled sequencer with DNBSEQ-G400 flow cell (see magnified inset). **d,** Pressure distribution in the DNBSEQ-G400 flow cell (rainbow colors); 170-µm-thick cover glass deformation (blue-to-red colors). Both by CFD, see Extended Data Fig. 1f-h for meshing details. **e**, Relative reagent concentration in the flow cell for five successive time (side views). Both cells in **(a-e)** have L = 7 cm, h = 20 µm, and V_a = 0.16 m/s. The width (w) for the r2r-fl cell is 7 cm, and 5.2 mm for the flow cell. **f-h**, Dependence of pressure drop average flow velocity (V_a), chamber length (L), and gap height (h), respectively. The pressure limit for the flow cell (e.g., microvalves or gasket) is 90 kPa (dashed line). **i, j**, Dependence of exchange ratio on average flow velocity **(i)** and chamber length **(j)**. The exchange ratio is defined as the supplied reagent volume required to achieve 99.9% purity at the biochip surface divided by the chamber volume. **k,** Dependence of the number of dNTP molecules supplied to each nanowell, on the tubing volume relative to the chamber volume for the flow cell (amplification N = 250 for DNA nanoballs). The dashed line shows the exchange ratio in r2r-fl, which does not require tubing. See also Supporting Video 1.

Figure 3 | Biochemical aspects of NGS using r2r-fl. a, Scheme showing how five sequential reagent/washing bands traverse the biochip, each reagent is deposited onto the moving PET belt by a dedicated slot die. Order is HOT, COLD, WB2, RR, RB. See Supporting Movie 2. **b,** Side view schematic of an operational unit for a single biochip, where the chip-holder increases to 57 °C when using the "HOT" sequencing reagent. **c,** Workflow for preparation of the DNA nanoballs (DNBs). Single-stranded DNA fragments are extracted from target DNA and amplified 250 times through rolling circle amplification. Each fragment forms a nanoball. **d,** Selective attachment of DNBs to the patterned biochip through ionic bonding. **e,** Schematic of the fluorescent imaging process for base calling. **f,** BGI's SBS biochemistry to sequence a DNB. The sequencing reagent (HOT) contains fluorescent dNTP (N stands for A, T, C or G) molecules and performs a self-terminated reaction to extend one complementary base. The sequencing reagent (COLD) reaction contains non-fluorescent dNTP to extend one self-

terminated complementary base with reaction points that failed to react with HOT. The biochip is then washed with WB2 and base calling is done by fluorescence imaging. Regeneration Reagent (RR) removes the fluorescent and azide groups to enable the next synthesis reaction cycle and is washed by Regeneration Buffer (RB). **g,** After flushing with washing buffer 2 (WB2), four-channel fluorescent imaging detects which type of nucleotide is added to each DNB, as shown in the merged image and the inset. **h,** Picture of the smallest 1×1 cm^2 and largest 15×15 cm^2 sequencing biochip (of the current work), and a single channel 6.3×0.75 cm^2 flow cell (DNBSEQ-200). **i**, Bar chart performance comparison of flow cell (DNBSEQ T7 and Novaseq X) versus r2r-fl for speed, reagent consumption, and data size. TAT is turnaround time for PE 100.

Figure 4 | Data quality comparison of next-generation sequencing (NGS) using r2r-fl versus flow cells. a-c, Typical heatmaps for the proportion of calling one incorrect base in 1000 bases (Q30) for r2r-fl, BGI DNBSEQ-T7, and Illumina NovaSeq X flow cells. Each dot represents a square field of view (FOV) for microscopy and is color-coded based on Q30, which is the benchmark for NGS. In (**a**), a selected region used for image registration is enclosed in a circle. In (**b**, **c**)**,** Q30 distribution indicates non-uniformity at edges due to uneven reagent distribution. White spots in (**b**) represent glue posts supporting the cover glass to prevent deformation. **d**, Fluorescence intensity of all four-color channels for r2r-fl, and two channels (green and blue) with different combinations to indicate the four nucleotides for Illumina NovaSeq X. **e,** Lag dependency on cycle number, representing the fraction of reaction points on DNBs falling behind the current cycle number. **f,** Run on dependency on cycle number, indicating the fraction of reaction points on DNBs exceeding the current cycle number. **g,** Base distribution for each sequencing cycle. **h,** Fraction of reads exceeding a base call accuracy of 99.9% (Q30 or higher). **i,** Precision and sensitivity for single nucleotide polymorphism (SNP) variants and insertions and deletions (INDEL) for r2r-fl and flow cell. Panels **d-i**, flow cell means the NovaSeq X. Illumina data is sourced from https://basespace.illumina.com.

Reference

1. Margulies, M. et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380 (2005).
2. Holt, R.A. & Jones, S.J. The new paradigm of flow cell sequencing. *Genome Res.* **18**, 839-846 (2008).
3. Uhlen, M. & Quake, S.R. Sequential sequencing by synthesis and the next-generation sequencing revolution. *Trends Biotechnol.* (2023).
4. Lu, H., Stratton, C.W. & Tang, Y.W. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *J. Med. Virol.* **92**, 401 (2020).
5. Kozlov, M. Ebola outbreak in Uganda: how worried are researchers. *Nature* **597** (2022).
6. Tegally, H. et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438-443 (2021).
7. Keita, A.K. et al. Resurgence of Ebola virus in 2021 in Guinea suggests a new paradigm for outbreaks. *Nature* **597**, 539-543 (2021).
8. Dudas, G. et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**, 309-315 (2017).
9. Fuller, C.W. et al. The challenges of sequencing by synthesis. *Nat. Biotechnol.* **27**, 1013-1023 (2009).
10. Balasubramanian, S. Sequencing nucleic acids: from chemistry to medicine. *Chem. Commun.* **47**, 7281-7286 (2011).
11. Rodriguez, R. & Krishnan, Y. The chemistry of next-generation sequencing. *Nat. Biotechnol.* (2023).
12. Case, D.J., Liu, Y., Kiss, I.Z., Angilella, J.-R. & Motter, A.E. Braess's paradox and programmable behaviour in microfluidic networks. *Nature* **574**, 647-652 (2019).
13. Bruus, H. Theoretical Microfluidics. (Oxford University Press, 2007).
14. Eisenstein, M. Innovative technologies crowd the short-read sequencing market. *Nature* **614**, 798-800 (2023).
15. Mohammed, M.I., Haswell, S. & Gibson, I. Lab-on-a-chip or Chip-in-a-lab: Challenges of Commercialization Lost in Translation. *Procedia Technology* **20**, 54-59 (2015).
16. Schwarze, K. et al. The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genet. Med.* **22**, 85-94 (2020).
17. Chai, J.H. et al. Cost-benefit analysis of introducing next-generation sequencing (metagenomic) pathogen testing in the setting of pyrexia of unknown origin. *PLoS One* **13**, e0194648 (2018).
18. Schermelleh, L. et al. Super-resolution microscopy demystified. *Nat. Cell Biol.* **21**, 72-84 (2019).
19. Heintzmann, R. & Huser, T. Super-resolution structured illumination microscopy. *Chem. Rev.* **117**, 13890-13908 (2017).
20. Ma, W. et al. Gene sequencing reaction device, gene sequencing system, and gene sequencing reaction method. *US11857973B2* (2024).
21. Beckett, N. & Caswell, N. Implementing barriers for controlled environments during sample processing and detection. *US20210354126A1* (2022).
22. Barbee, K. et al. Methods for biological sample processing and analysis. *US11499962B2* (2019).
23. Pennisi, E. $100 genome? New DNA sequencers could be a 'game changer' for biology, medicine. *Science*.
24. Reynolds, O. IV. On the theory of lubrication and its application to Mr. Beauchamp tower's experiments, including an experimental determination of the viscosity of olive oil. *Philos. Trans. R. Soc. London*, 157-234 (1886).
25. Reichardt, H. Über die Umströmung zylindrischer Körper in einer geradlinigen Couetteströmung. *(No Title)* (1954).

26. Tillmark, N. & Alfredsson, P.H. Experiments on transition in plane Couette flow. *J. Fluid Mech.* **235**, 89-102 (1992).
27. Krug, D., Lüthi, B., Seybold, H., Holzner, M. & Tsinober, A. 3D-PTV measurements in a plane Couette flow. *Exp. Fluids* **52**, 1349-1360 (2012).
28. Piau, J., Bremond, M., Couette, J. & Piau, M. Maurice Couette, one of the founders of rheology. *Rheol. Acta* **33**, 357-368 (1994).
29. Fardin, M., Perge, C. & Taberlet, N. "The hydrogen atom of fluid dynamics"–introduction to the Taylor–Couette flow for soft matter scientists. *Soft Matter* **10**, 3523-3535 (2014).
30. Mardis, E.R. DNA sequencing technologies: 2006–2016. *Nature protocols* **12**, 213-218 (2017).
31. Rodriguez, R. & Krishnan, Y. The chemistry of next-generation sequencing. *Nat. Biotechnol.* **41**, 1709-1715 (2023).
32. Ali, M.M. et al. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. *Chem. Soc. Rev.* **43**, 3324-3341 (2014).
33. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).
34. Krusche, P. et al. Best practices for benchmarking germline small-variant calls in human genomes. *Nat. Biotechnol.* **37**, 555-560 (2019).
35. Xiao, K. et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* **583**, 286-289 (2020).
36. Foox, J. et al. Performance assessment of DNA sequencing platforms in the ABRF Next-Generation Sequencing Study. *Nat. Biotechnol.* **39**, 1129-1140 (2021).
37. Zheng, W. et al. High-throughput, single-microbe genomics with strain resolution, applied to a human gut microbiome. *Science* **376**, eabm1483 (2022).
38. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865-868 (2017).
39. Lewis, S.M. et al. Spatial omics and multiplexed imaging to explore cancer biology. *Nat. Methods* **18**, 997-1012 (2021).
40. Bressan, D., Battistoni, G. & Hannon, G.J. The dawn of spatial omics. *Science* **381**, eabq4964 (2023).
41. Chiu, C.Y. & Miller, S.A. Clinical metagenomics. *Nature Reviews Genetics* **20**, 341-355 (2019).
42. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *New Engl. J. Med.* **382**, 727-733 (2020).
43. Church, A.J. et al. Molecular profiling identifies targeted therapy opportunities in pediatric solid cancer. *Nat. Med.* **28**, 1581-1589 (2022).
44. Nogrady, B. How cancer genomics is transforming diagnosis and treatment. *Nature* **579**, S10-S11 (2020).
45. Saunders, C.J. et al. Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. *Science translational medicine* **4**, 154ra135-154ra135 (2012).
46. Sazonovs, A. et al. Large-scale sequencing identifies multiple genes and rare variants associated with Crohn's disease susceptibility. *Nat. Genet.* **54**, 1275-1283 (2022).

Methods

Genomic materials. The genomic samples used in this study were obtained from the following sources: i) *E. coli* samples were derived from the MGI Standard Library Kit V3.0 (Part Number 1000005033); ii) human genome samples were obtained from the NIGMS Human Genetic Cell Repository (NA12878) by Coriell Institute, using the MGIEasy PCR-Free DNA Library Prep Set V1.1 (16 RXN, Part Numbers 1000013452); iii) SARS-CoV-2 genome samples were sourced from the SARS-CoV-2 Molecular Controls Kit - Full Genome (Part Numbers GW-CRBM002) by GeneWell Biotechnology Co., Ltd; and MGI ATOPlex RNA Multiplex PCR-based Library Preparation Set V3.0 (16 RXN, Part Numbers 940000132-00).

CFD simulations. COMSOL 6.0 was utilized for conducting fluid simulations in our study. We assumed standard Newtonian fluids with a viscosity of μ=10^{-3} Pa·s, and the diffusion coefficient of biomolecules was set to D=10^{-9} m^2/s. The mass transfer and biomolecule diffusion phenomena were calculated using the Navier-Strokes equation $\rho(\partial_t \boldsymbol{V} + (\boldsymbol{V}\cdot\nabla)\boldsymbol{v}) = -\nabla p + \mu\nabla^2\boldsymbol{V}$ and the convection-diffusion equation $\partial_t C + \boldsymbol{V}\cdot\nabla C = D\nabla^2 C$. To calculate the deformation of the cover glass, we applied the equations of motion $0 = \nabla\cdot(\mathrm{FS})^T + F_v$ and $F = \mathrm{I} + \nabla u_{solid}$. In a typical experiment, the biochip size was set to 7 cm × 7 cm for roll-to-roll fluidics (r2r-fl). The belt velocity was set to V_s = 0.32 m/s, resulting in an average flow velocity V_a = 0.16 m/s, and the gap height h = 20 μm. The model flow cell (see Extended Data Fig. 1f–h) was configured with a 30-cm inlet tubing, and the dimensions of the biochip were 72.3 mm in length and 5.2 mm in width.

We used standard quadrilateral meshing for our simulations. The flow cell chamber was configured with predefined normal settings calibrated for fluid dynamics. The inlet and outlet tubing had free quad bottoms, with 50 and 20 layers respectively, and element ratios of 100 and 20, symmetrically distributed. The biochip had 8 layers with a free quad bottom set to "finer" and calibrated for fluid dynamics. The outer surfaces of the tubing edges of the biochip had a boundary layer with 4 layers. For the r2r-fl, we used the following mesh size: the upper surface consisted of free quadrilateral nodes with a maximum size of 1 cm, swept for 6 layers with an element ratio of 6 and symmetric distribution.

Fabrication of biochips. Biochips comprising arrays of 2.5 × 10^{10} nanowells were fabricated from 7 cm × 7 cm silicon wafers utilizing deep ultraviolet lithography. In this process, hexamethyldisilane (HMDS) was vaporized and carefully deposited on the oxidized silicon substrate. HMDS formed a robust bond with OH groups on the silicon surface, resulting in the formation of a hydrophobic surface due to the attachment of methyl groups. These methyl groups remained on the surface, ensuring strong adhesion with the photoresist material. To achieve a desired thickness of 500 nm, the photoresist was spin-coated onto the wafers. Employing deep ultraviolet (DUV) lithography, a 15 × 15 cm^2 array of 10 × 10^{10} circular wells with a diameter of approximately 250 nm and a center-to-center separation of 500 nm was precisely created to enable one-to-one attachment of DNA nanoballs (DNBs). To eliminate any remnants of the photoresist from the bottom of the nanowells, the biochip underwent a thorough washing procedure involving acetone, followed by oxygen dry etching. Furthermore, through the application of the chemical vapor deposition (CVD) technique and amination of the silicon surface of the biochips by exposing them to HMDS vapor, the methyl groups on the surface were effectively converted into amine groups. This transformation rendered the nanowells hydrophilic and positively charged, facilitating the attachment of DNBs. Following the removal of all photoresist material, the biochip exhibited hydrophobic properties, except for the nanowells. Subsequently, a laser was employed to cut the wafer into the desired design shape, without necessitating any additional polishing steps.

Finally, optional 20 µm thick spacers, composed of 10 µm 3M UCT-10 double-stick tape and a 10 µm stainless steel shim, were affixed to the edges of the biochip. Handles, necessary for manipulation by a robotic arm, were attached using 3M adhesives.

Preparation of DNA nanoballs (DNBs). DNA nanoballs (DNBs) were prepared for various genomes[47], including the SARS-CoV-2 virus genome, human genome (after the aforementioned extraction steps), and *E. coli* from the Standard Library Kit V3.0 (Part Number 1000005033). The preparation was performed using the BGISEQ-500RS DNB Make Load Reagent (Part Number 1000005488) from MGI Tech Co. Ltd. The following steps were conducted: Initially, DNA was annealed using a Bio-Rad S1000TM thermocycler with temperature cycles of 95°C for 1 min, followed by 65 °C for 1 min, and 40 °C for 1 min [48]. Subsequently, 40 µL of Make DNB Enzyme Mix I and 4 µL of Make DNB Enzyme Mix II were added for rolling circle amplification, which took place at 30°C for 20 min. To terminate the amplification, the temperature was lowered to 0°C, and 20 µL of Make DNB Stop Buffer was added. This procedure resulted in the generation of 250 faithful copies of the libraries that formed DNA nanoballs (DNBs) with a normal size of ~ 220 nm (ref. [58]) and a concentration of 15-20 ng/µL.

Loading biochips with DNBs. To facilitate the adhesion of DNA nanoballs (DNBs) onto the wells of the biochip, the biochip was carefully placed into a loader, which functioned as a flow chamber with an 80 µm gap height set at 25 °C. The DNB solution was prepared by combining 400 µL of DNB Reagent, 500 µL of DNB Load Buffer I, and 300 µL of DNB Load Buffer II. To ensure uniform coating of the biochip's surface, a total of 1480 µL of the DNB solution was prepared, taking into account a reagent loss of 410 µL as bypass volume from the previous section into the loader. Denatured-free phi29 polymerase was adsorbed onto the aminated spots of the biochip, while Pluronic F68 (PF68), a PEO-PPO-PEO tri-block copolymer, prevented DNBs from connecting to the hexamethyldisilazane (HMDS) field through steric hindrance. PF68 contains hydrophilic hydroxyl groups at both ends (PEO part) and hydrophobic methyl groups in the middle (PPO part). When PF68 dissolved in DNB Load Buffer II flowed onto the biochip surface along with the DNA nanoballs, the principle of "like dissolves like" came into play. Due to electrical interactions between polar molecules, non-polar molecules in polar solvents tend to aggregate to minimize the hydrophobic area. Consequently, the hydrophobic PPO part of PF68 is attracted to the methyl groups of HMDS, causing the negatively charged DNBs to attach singly to the positively charged nanowells. To remove excess DNBs, a rinse with DNB Rinse Buffer (DRB) was performed. Subsequently, DNBs were condensed using DNB Crash Buffer (DCB). The biochips were then flushed with DNB Read Buffer (DRB) and Protein Washing Buffer (PWB). DRB facilitated the binding between DNBs and the biochip, while bovine serum albumin (BSA) in PWB was adsorbed by DNBs to provide support. To enhance stability, the biochips were flushed with DCB and REB once again. After these steps, the biochips were treated with Hybridization Buffer (HYB) and soaked in 1 µM AD153 Sequencing Primer 1 V3.0 at different temperatures (55°C for 9 minutes, 40°C for 2 minutes, and 25°C for 2 minutes). Based on the principle of complementary base pairing, the primers were spliced onto the corresponding single-stranded fragments on the DNB beads, enabling subsequent sequencing. To evaluate the loading effect, an initial round of sequencing was performed on the loader. After the sequencing primers hybridized to the adapter region of the DNBs, a fluorescently labeled dNTP probe was incorporated using cPAS DNA polymerase. Any unbound dNTP probes were then washed away. The DNB Flow Cell was imaged, and the fluorescence signal was converted to a digital signal. The base information was determined using MGI's proprietary base-calling software. To prevent the DNB-loaded biochips from drying out, they were stored in Wash Buffer 2 (WB2). All

reagents used in these processes were sourced from the DNBSEQ-T10x4 RS High-Throughput Sequencing Reagent Kit (Part Number 1000005488 from MGI Tech Co. Ltd).

Reagents for biochemistry. For sequencing, we conducted the standard Massively Parallel Sequencing (MPS) process using the DNBSEQ-T10x4 RS High-Throughput Sequencing Reagent Kit (category number 1000023411, MGI Tech Co. Ltd) designed for FCL PE100 sequencing. The process involves three distinct biochemical reactions as follows: 1) to replicate the complementary bases of the DNB template, we utilized Sequencing Reagent (HOT), which consists of cPAS DNA polymerase and dNTP molecules (N = A, T, C, or G). The dNTP molecules are conjugated with fluorophores that can be spectrally resolved by the imaging system[15]; 2) for bases of the DNB template that did not react with Sequencing Reagent (HOT), we employed non-fluorescent Sequencing Reagent (COLD), comprising unlabeled dNTPs and cPAS DNA polymerase for complementary base replication; and 3) To remove the fluorescent cap in preparation for the next sequencing cycle, we utilized Regeneration Reagent 2. The kit provides specific buffer washing solutions for each of these three reagents, including Wash Buffer, Sequencing Reagent, and Regeneration Buffer. It is important to note that unlike conventional NGS, there are no concerns regarding cross-contamination, allowing the elimination of the buffer for Sequencing Reagent (HOT) and direct flow of non-fluorescent Sequencing Reagent (COLD) immediately after Sequencing Reagent (HOT). To prevent beading up on the biochip and PET belt, the surface tension of all six solutions was reduced to less than 42 mN/m using Tween-20 surfactant. The first strand sequencing occurs from cycle 1 to cycle 100. Upon completion of the 1st strand sequencing, the 2nd strand generation primers and a polymerase with strand displacement activity are introduced to initiate 2nd strand synthesis. The polymerase extends the new primer until it reaches the original sequenced strand, displacing the original sequencing strand and forming a new single-stranded template. To maximize the length of the strand while ensuring its attachment to the original DNB, the newly generated 2nd strand is optimized. Following hybridization of the 2nd strand sequencing primer through multiple displacement amplification (MDA), the same sequencing chemistry employed during 1st strand sequencing is used for 2nd strand sequencing. The newly generated 2nd strand template contains a higher number of copies of the insert DNA, resulting in stronger signals and increased sequencing accuracy.

Roll-to-roll fluidics (r2r-fl) platform setup. To establish r2r-fl, a custom slot-die coater and the biochip holder from DNBSEQ-T7 were integrated onto an industrial roller (Siemens 6SL3210-5FE10-4UA0). The setup included five slot-die coating units (Foshan Edge Development Mold Mechanical Technology, Inc) corresponding to the five reagent formulations discussed previously. The coating roller served as the driving mechanism, while the tension of the PET belt was carefully adjusted to 20 kg to ensure flatness. The belt speed was set to 0.32 m/s, and the slot-die flow rate was meticulously adjusted to achieve a 20-μm gap thickness. To render the belt hydrophilic, a corona pretreatment was employed. The PET belt had a thickness of 100 μm and a total length of 3000 m, allowing for up to 500 sequencing cycles.

Roll-to-roll fluidics platform operation. A Robotic Arm, C8XL by Epson, facilitated the movement of the biochip between the r2r-fl platform and the imaging subsystems, which was adapted from the commercial DNBSEQ-T10 by BGI. During the intervals between processing cycles, the biochip was securely stored in a dedicated transfer slot containing Wash Buffer 2. The biochemical processes involved in base calling comprised two steps for base addition and one step for fluorescent group cleavage, necessitating three reagent bands and two washing bands. Prior to operation, the slot dies were primed to ensure optimal performance. Each band had a length of 40 cm, which was more than twice the size of the biochip, ensuring complete replacement of the reagents. A complete biochemical cycle

was accomplished within 60 seconds. The transfer of the biochip from the transfer slot bath to the imaging subsystem and back, facilitated by the robotic arm, was completed in 20 seconds. The imaging process itself required 60 seconds.

Reagent replacement in the flow cell and r2r-fl. The cross-contamination of dNTP synthesis reagents (HOT and COLD in BGI's SBS chemistry) and the regeneration reagent, which removed the steric hindrance group, should be avoided as much as possible. Otherwise, more than one base would be added to each reaction site and drive run on surge. In quality control standard, it is less than 0.1% mixing. In practice, we utilized two reagents (one of them was a buffer) to achieve this goal. To evaluate the cleaning efficiency for the flow cell and r2r-fl cell, we employed the standard method in R&D to compare the fluorescent replacement speed of the actual sequencing reagent (HOT). First, we filled a new flow cell or a r2r-fl cell with Wash Buffer 2, took pictures near the outlet using the imaging module of the sequencer, and considered the average gray value before any fluorescent reagent contamination as the background value. Next, we diluted HOT about twice until the average gray value was close to but less than the saturation value of 65535, considering it as the full fluorescent value. Then, we used Wash Buffer 2 to clean the HOT until the remaining gray value was less than "the background value + 0.1% x (full fluorescent value - background value)". For example, if the full fluorescent value of the flow cell is 63000, and its background value is 300, then its endpoint was considered the moment of gray value for the outlet reaching "300 + (63000-300) = 362.7". We used DNBSEQ-G50 and its flow cell with a 7 cm length and 50 μm gap height, along with a 7 × 7 cm^2 r2r-fl cell with 50 μm gap height for comparison. The flow cell was connected to the Wash buffer 2 bottle with a 30 cm tubing with a 0.8 mm inner diameter, shorter than the actual product of DNBSEQ-G50. For the gray value at a specific moment (e.g., 20 s) in the figure, we run the pump/PET belt for a certain time and stopped, took a picture using the imaging module, and recorded the average gray value close to the outlet on the biochip. Each data point was repeated at least three times.

Reagent layer uniformity and thickness. We achieved uniform coatings by carefully adjusting the surface energy of both the reagents and the polymeric belt. To promote spreading and prevent beading up, we employed corona pretreatment to increase the PET belt (Nanya plastic corporation, Rynite 935 BK505 with 35% mica/glass reinforced modified polyethylene terephthalate resin) surface energy. Additionally, we incorporated 0.5% Tween-20 to reduce the surface tension of the reagents. The thickness of the layers was measured using the Keyence sensor SI-T80, which offers a linearity of ±0.05% of full scale and a repeatability of 0.5 μm. To determine the actual liquid thickness, we moved the laser detector across the wet zone of the PET belt from one end to the other, resulting in a measured thickness of 10±2 μm with a variation of ± 0.5 μm.

Identification of COVID-19 variants. We conducted a comprehensive analysis to ensure the traceability of COVID-19 data on our platform, including the identification of variants. The COVID-19 sequence data obtained from our platform was aligned to the SARS-CoV-2 wild-type reference genome (Genbank: MN908947) using the BWA-MEM alignment tool. Variants in the alignment results were called using freebayes, and the results were further annotated using SnpEff for comprehensive variation calling. To validate the accuracy of the called variants, we compared them with the reference genome and generated a FASTA file containing the sequenced library with the identified variants. Notably, our analysis revealed no variants, indicating that the sequence of the library perfectly matched the reference genome. To further confirm the traceability, we utilized Nextclade [49] (https://clades.nextstrain.org) to track the FASTA file of the sequenced library. The traceable results demonstrated that the sequence

library exhibited a high degree of homology with the SARS-CoV-2 wild-type reference genome (Genbank: MN908947).

Whole genome sequencing (WGS) analysis. The data were processed using SOAPnuke [50] and aligned to the human reference genome (hs37d5) using the BWA-MEM [51] alignment algorithm. SAMtools was then utilized to sort the alignment results [52], and Picard was employed to mark duplicates after the sorting process [53]. GATK3.7 was applied to correct the base quality values using the BQSR (Base Quality Score Recalibration) algorithm, following the parameter recommendations provided by the official GATK website [54]. Germline mutations were called using GATK HaplotypeCaller. Copy number variations were analyzed using CNVnator [55], while structural variants were assessed using BreakDancer (V1.0) [56]. To provide comprehensive annotations, the identified variants were annotated using ANNOVAR and RTG vcfeval. These rigorous data processing and analysis steps utilizing established tools and algorithms ensure the accuracy and reliability of the obtained results.

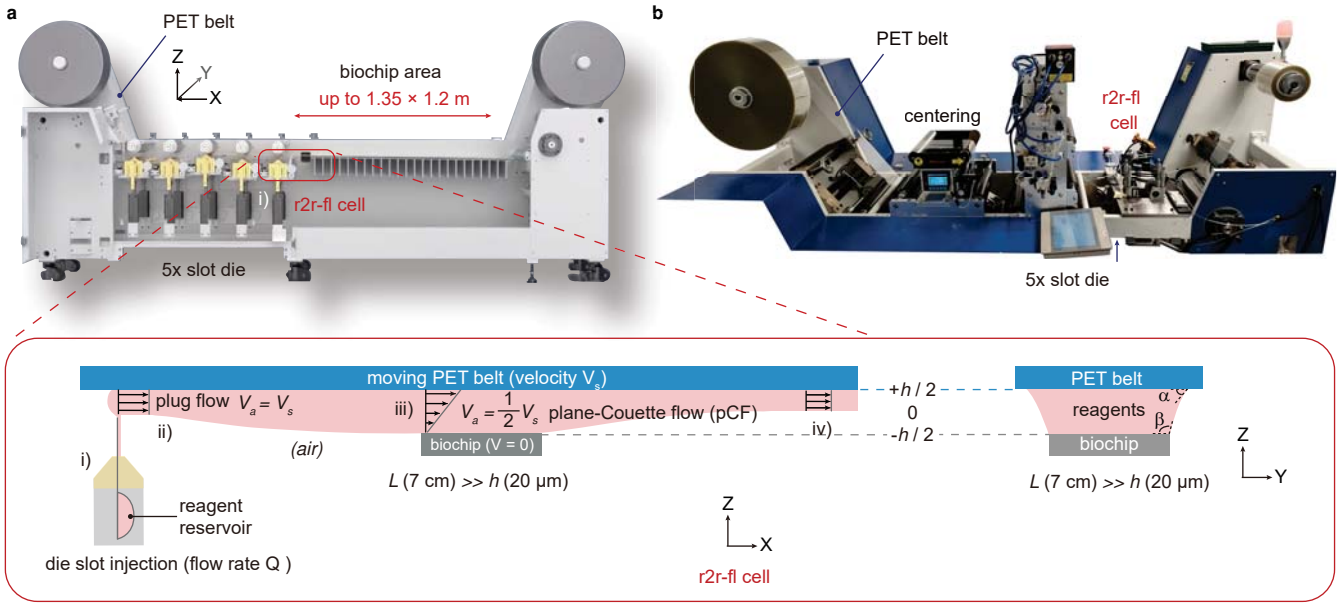
Data quality. Lag and run-on are attributed to imperfections in the biochemical synthesis procedure. Lag indicates that one or more bases have failed to be synthesized during the current sequencing cycle, while run-on signifies premature synthesis of one or more bases. To identify instances of lag and run-on, we utilized BGI's commercial software, which also generates informative figures and provides various indicators such as Q30. In order to determine the success of the loading process, we evaluated the Q30 report.

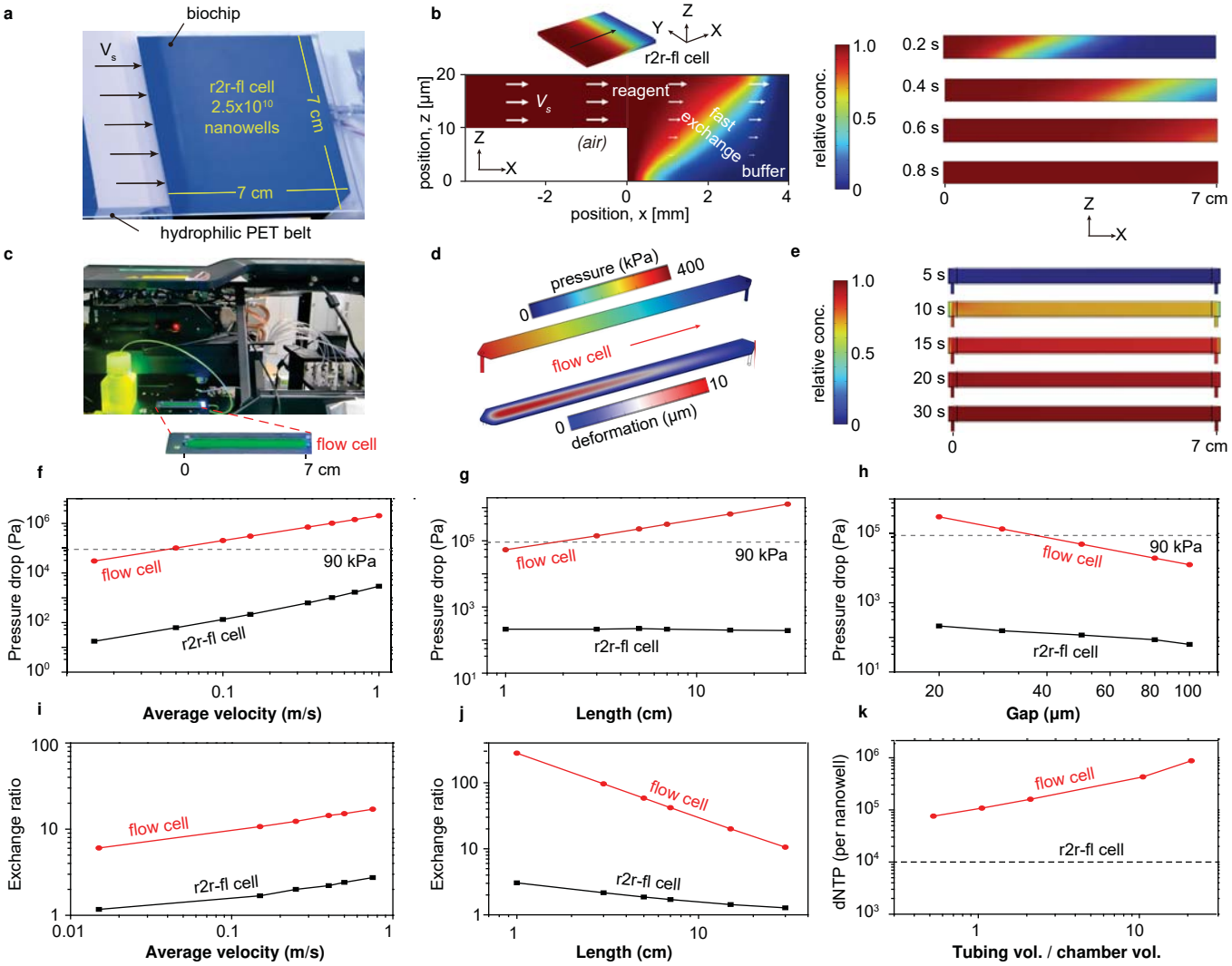
Data availability. The Illumina sequencing quality data used in this work for comparison are publicly available through basespace (basespace.illumina.com, registration needed), sequencing data and analysis are available in "Yanzhe Qin. Fast, cost-effective, and flexible DNA sequencing by roll-to-roll fluidics [DS/OL]. V1. Science Data Bank, 1[2024-12-24]. https://doi.org/10.57760/sciencedb.18851. DOI:10.57760/sciencedb.18851." Other data that support the findings of this study are available from the corresponding author upon reasonable request.

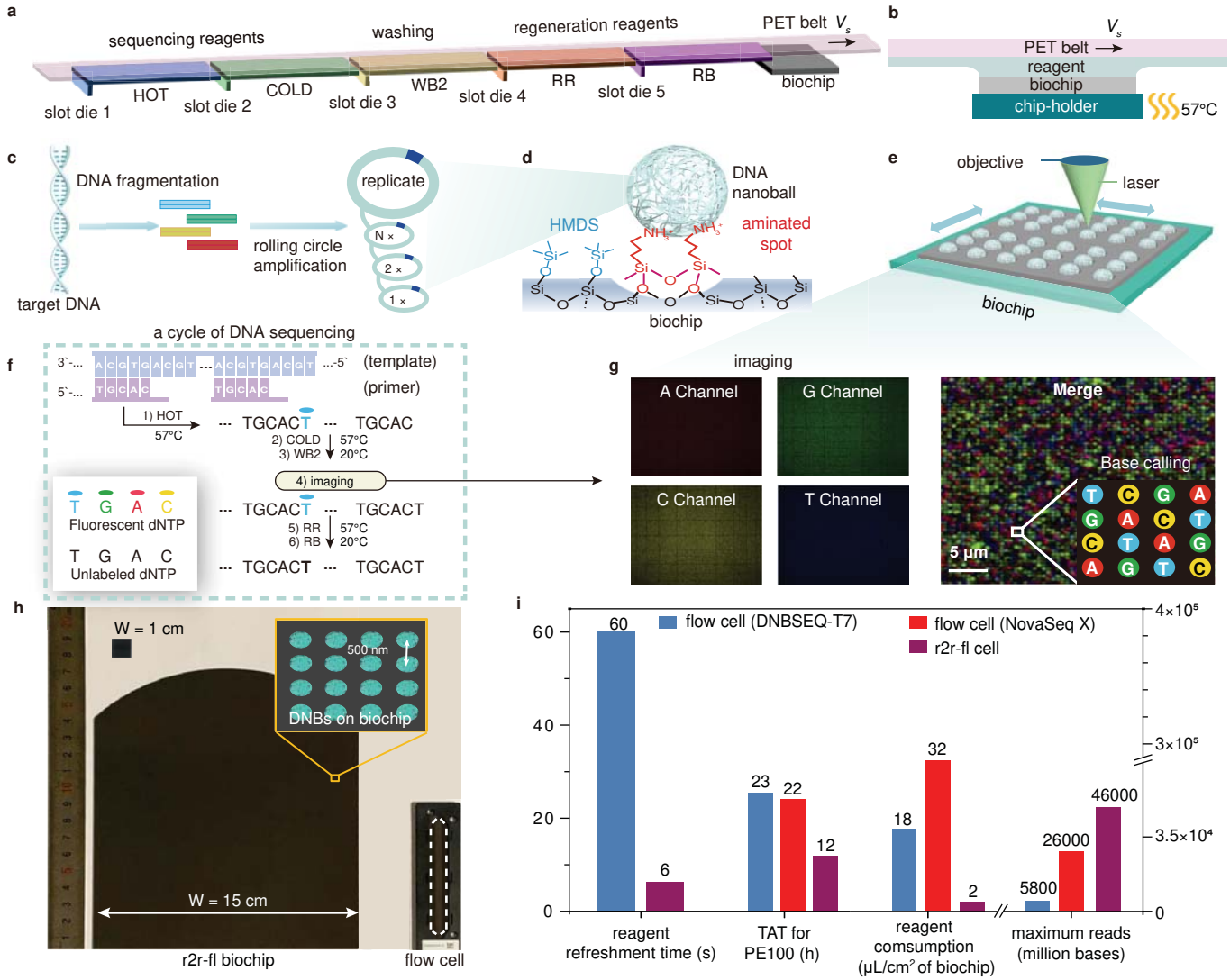
Method References

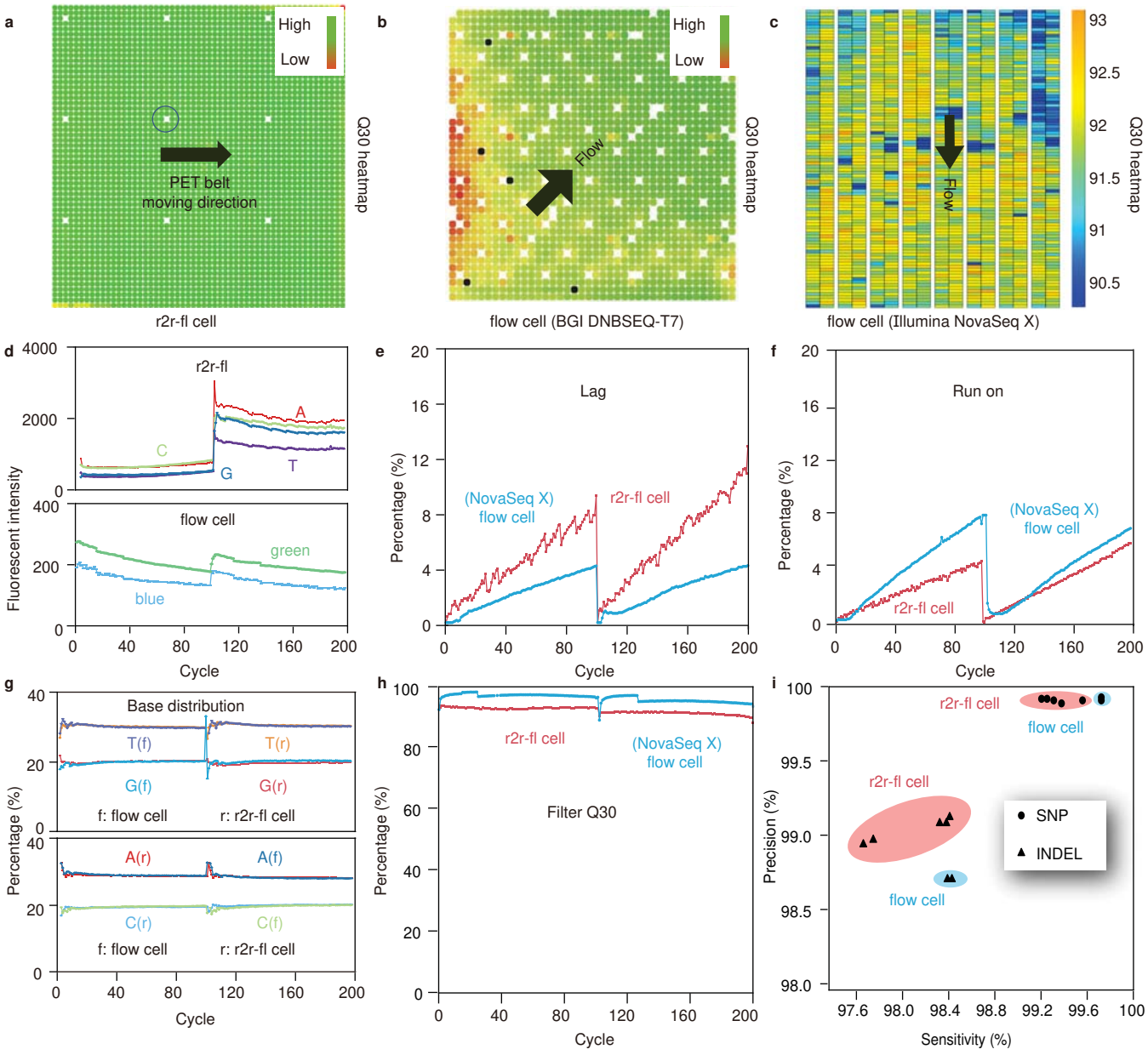
47. Drmanac, R. et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* **327**, 78-81 (2010).
48. Li, Z. et al. DNB-based on-chip motif finding: A high-throughput method to profile different types of protein-DNA interactions. *Science advances* **6**, eabb3350 (2020).
49. Aksamentov, I., Roemer, C., Hodcroft, E.B. & Neher, R.A. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *Journal of Open Source Software* **6**, 3773 (2021).
50. Chen, Y. et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* **7**, gix120 (2018).
51. Jung, Y. & Han, D. BWA-MEME: BWA-MEM emulated with a machine learning approach. *bioRxiv* (2021).
52. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
53. Toolkit, P. (2019).
54. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*, 201178 (2017).
55. Abyzov, A., Urban, A.E., Snyder, M. & Gerstein, M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974-984 (2011).
56. Chen, K. et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* **6**, 677-681 (2009).

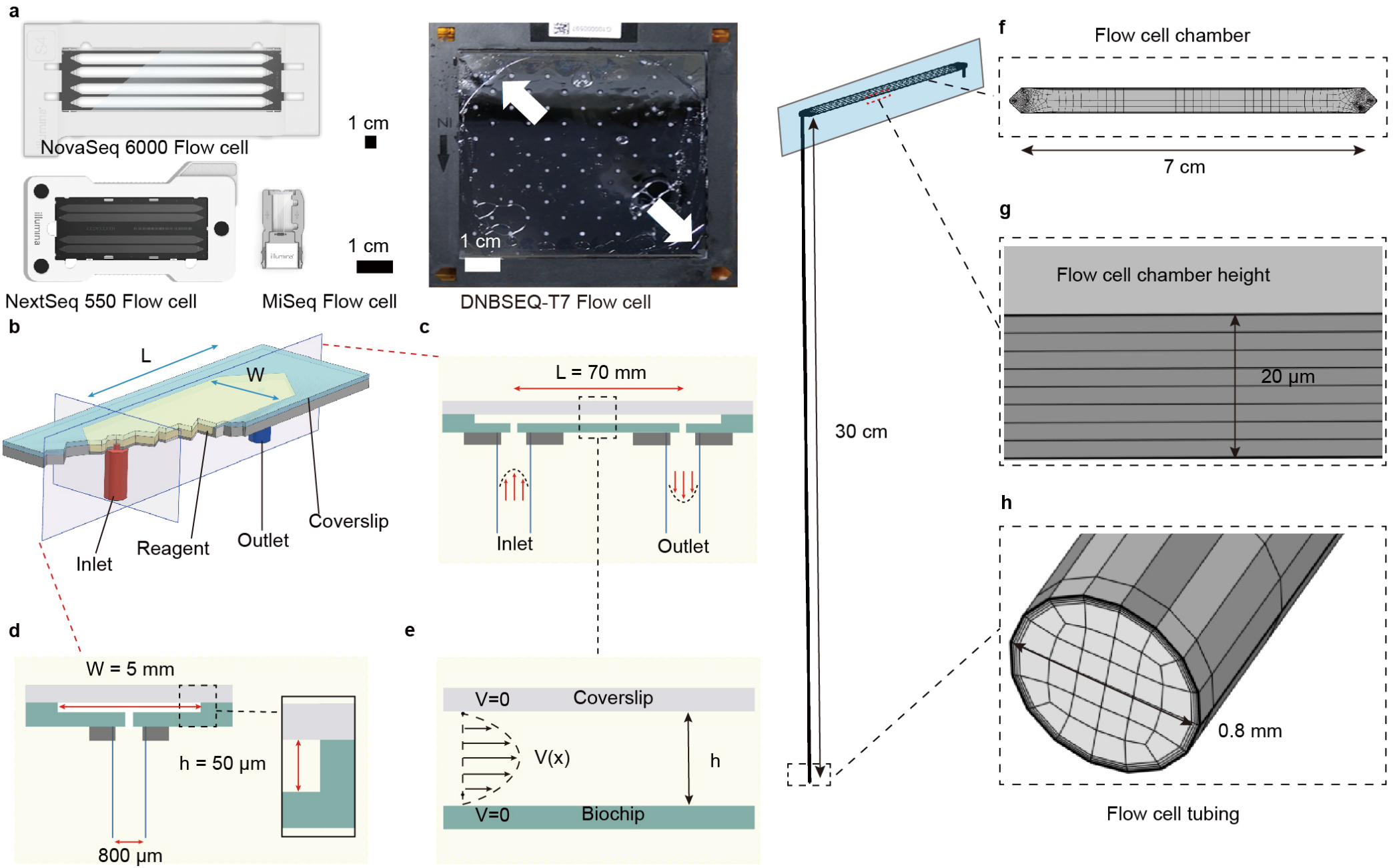
624


a
PET belt
Z
Y
X
biochip area
up to 1.35 × 1.2 m
i)
r2r-fl cell
5x slot die
b
PET belt
centering
r2r-fl
cell
5x slot die
moving PET belt (velocity V_s)
plug flow $V_a = V_s$
ii)
(air)
i)
reagent
reservoir
die slot injection (flow rate Q)
iii)
$V_a = \frac{1}{2} V_s$ plane-Couette flow (pCF)
biochip (V = 0)
L (7 cm) >> h (20 µm)
iv)
+h / 2
0
-h / 2
PET belt
reagents
α
β
biochip
L (7 cm) >> h (20 µm)
Z
Y
Z
X
r2r-fl cell
c
flow cell
r2r-fl cell
0
7 cm
d
volume (µL/cm²)
0
30
60
1
0
relative concentration*
99.9%
r2r-fl cell
flow cell
0
20
40
time [s]
e
price per DNA megabase
10,000
100
1
0.01
0.0001
US $
Moore's law
operational costs
reagents 89.3%
reagents
biochip (5.2%)
packaging (3.3%)
other (2.2%)
This
work
2000
2005
2010
2015
2020
2025
year

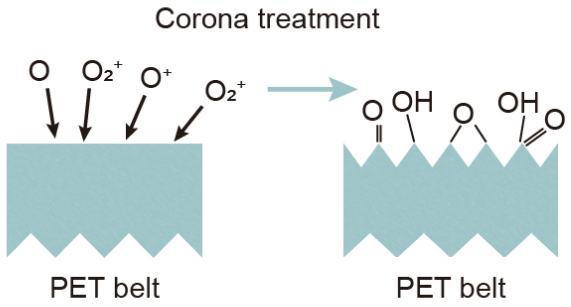

a
biochip
V_s
r2r-fl cell
2.5x10¹⁰
nanowells
7 cm
7 cm
hydrophilic PET belt
b
Z
Y
X
r2r-fl cell
position, z [µm]
20
10
0
V_s
reagent
(air)
fast
exchange
buffer
Z
X
-2
0
2
4
position, x [mm]
relative conc.
1.0
0.5
0
0.2 s
0.4 s
0.6 s
0.8 s
0
7 cm
Z
X
c
flow cell
0
7 cm
d
pressure (kPa)
400
0
flow cell
10
0
deformation (µm)
e
relative conc.
1.0
0.5
0
5 s
10 s
15 s
20 s
30 s
0
7 cm
f
Pressure drop (Pa)
10^6
10^4
10^2
10^0
flow cell
90 kPa
r2r-fl cell
0.1
1
Average velocity (m/s)
g
Pressure drop (Pa)
10^5
10^3
10^1
flow cell
90 kPa
r2r-fl cell
1
10
Length (cm)
h
Pressure drop (Pa)
10^5
10^3
10^1
flow cell
90 kPa
r2r-fl cell
20
40
60
80
100
Gap (µm)
i
Exchange ratio
100
10
1
flow cell
r2r-fl cell
0.01
0.1
1
Average velocity (m/s)
j
Exchange ratio
100
10
1
flow cell
r2r-fl cell
1
10
Length (cm)
k
dNTP (per nanowell)
10^6
10^5
10^4
10^3
flow cell
r2r-fl cell
1
10
Tubing vol. / chamber vol.


a
sequencing reagents
washing
regeneration reagents
PET belt V_s
slot die 1
HOT
slot die 2
COLD
slot die 3
WB2
slot die 4
RR
slot die 5
RB
biochip
b
V_s
PET belt
reagent
biochip
chip-holder
57°C
c
DNA fragmentation
target DNA
rolling circle amplification
replicate
N ×
2 ×
1 ×
d
DNA nanoball
HMDS
aminated spot
biochip
e
objective
laser
biochip
a cycle of DNA sequencing
f
(template)
(primer)
1) HOT 57°C
TGCACT ··· TGCAC
2) COLD 57°C
3) WB2 20°C
4) imaging
T G A C
Fluorescent dNTP
T G A C
Unlabeled dNTP
TGCACT ··· TGCACT
5) RR 57°C
6) RB 20°C
TGCACT ··· TGCACT
g
imaging
A Channel
G Channel
C Channel
T Channel
Merge
Base calling
5 μm
h
W = 1 cm
500 nm
DNBs on biochip
W = 15 cm
r2r-fl biochip
flow cell
i
flow cell (DNBSEQ-T7)
flow cell (NovaSeq X)
r2r-fl cell
60
40
20
0
4×10⁵
3×10⁵
3.5×10⁴
0
60
6
23
22
12
18
32
2
5800
26000
46000
reagent refreshment time (s)
TAT for PE100 (h)
reagent comsumption (μL/cm² of biochip)
maximum reads (million bases)

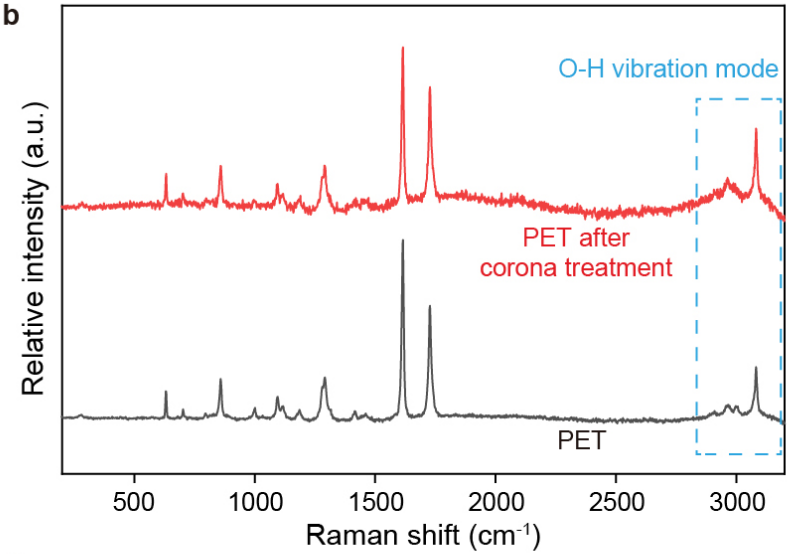

a
High
Low
PET belt
moving direction
Q30 heatmap
r2r-fl cell
b
High
Low
Flow
Q30 heatmap
flow cell (BGI DNBSEQ-T7)
c
93
92.5
92
91.5
91
90.5
Flow
Q30 heatmap
flow cell (Illumina NovaSeq X)
d
r2r-fl
A
C
G
T
flow cell
green
blue
Fluorescent intensity
Cycle
e
Lag
(NovaSeq X)
flow cell
r2r-fl cell
Percentage (%)
Cycle
f
Run on
(NovaSeq X)
flow cell
r2r-fl cell
Percentage (%)
Cycle
g
Base distribution
T(f)
T(r)
G(f)
G(r)
f: flow cell
r: r2r-fl cell
A(r)
A(f)
C(r)
C(f)
f: flow cell
r: r2r-fl cell
Percentage (%)
Cycle
h
r2r-fl cell
(NovaSeq X)
flow cell
Filter Q30
Percentage (%)
Cycle
i
r2r-fl cell
flow cell
r2r-fl cell
flow cell
SNP
INDEL
Precision (%)
Sensitivity (%)


a
NovaSeq 6000 Flow cell
1 cm
NextSeq 550 Flow cell
MiSeq Flow cell
1 cm
1 cm
DNBSEQ-T7 Flow cell
b
L
W
Inlet
Reagent
Outlet
Coverslip
c
L = 70 mm
Inlet
Outlet
d
W = 5 mm
h = 50 µm
800 µm
e
V=0
Coverslip
V(x)
h
V=0
Biochip
30 cm
f
Flow cell chamber
7 cm
g
Flow cell chamber height
20 µm
h
0.8 mm
Flow cell tubing

a



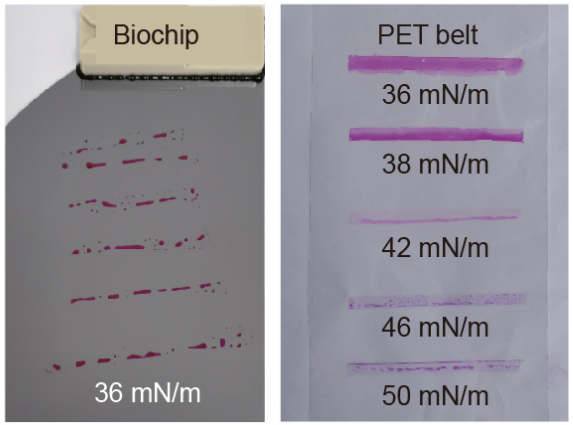
b



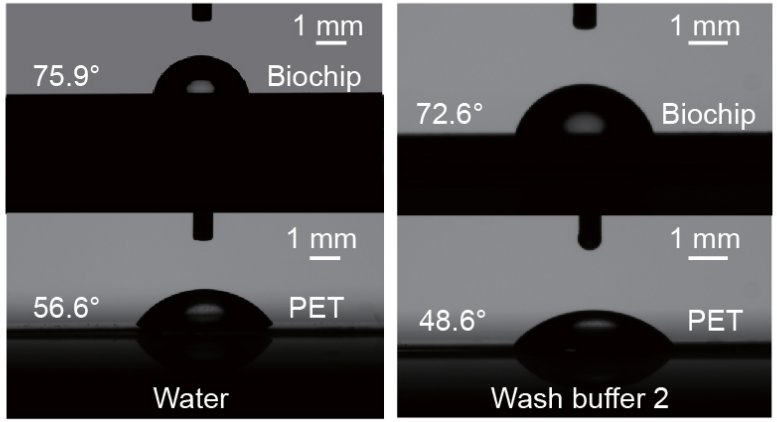
c

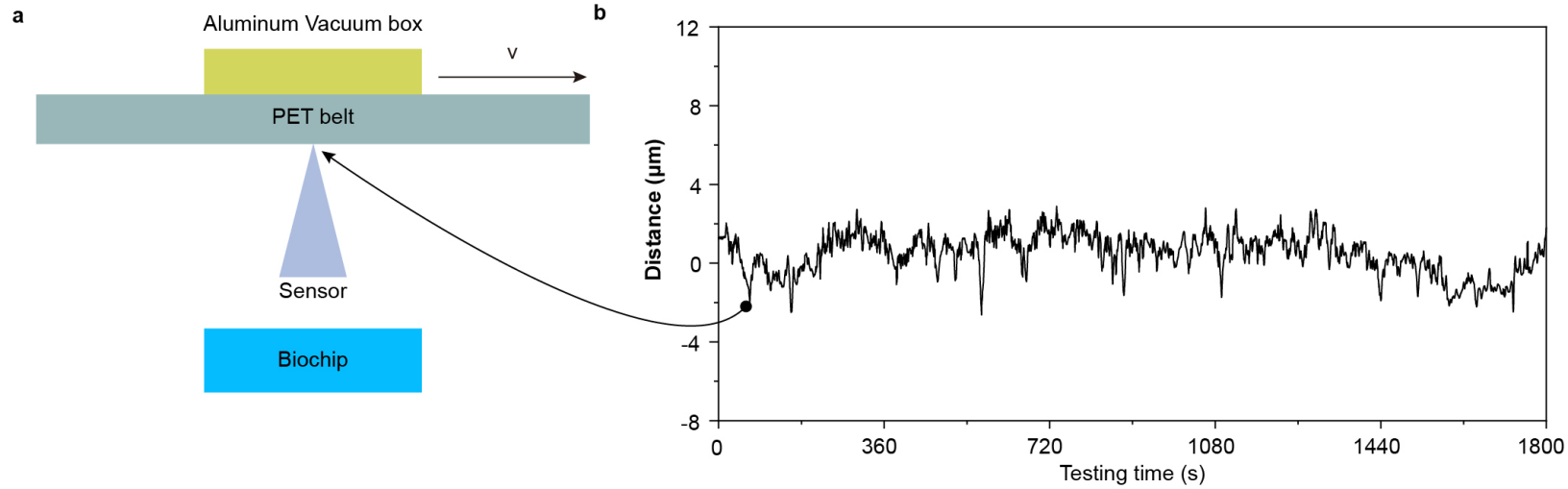
Reagent	Surface tension (mN/m)
Regeneration reagent	35.4
Regeneration buffer	38.8
Sequencing reagent (Hot)	32.3
Sequencing (Cold)	37.8
Wash buffer2	38.6

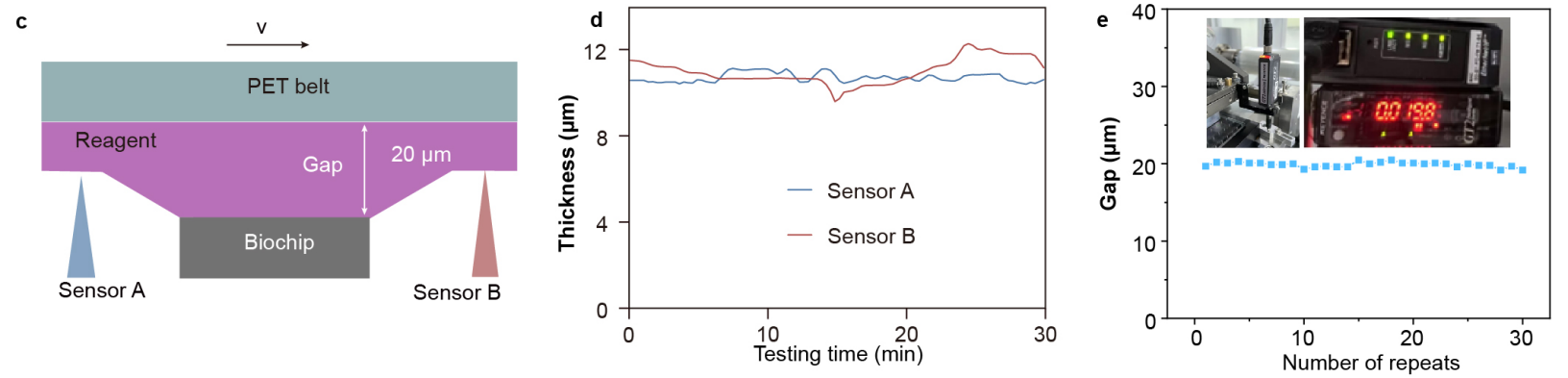
d

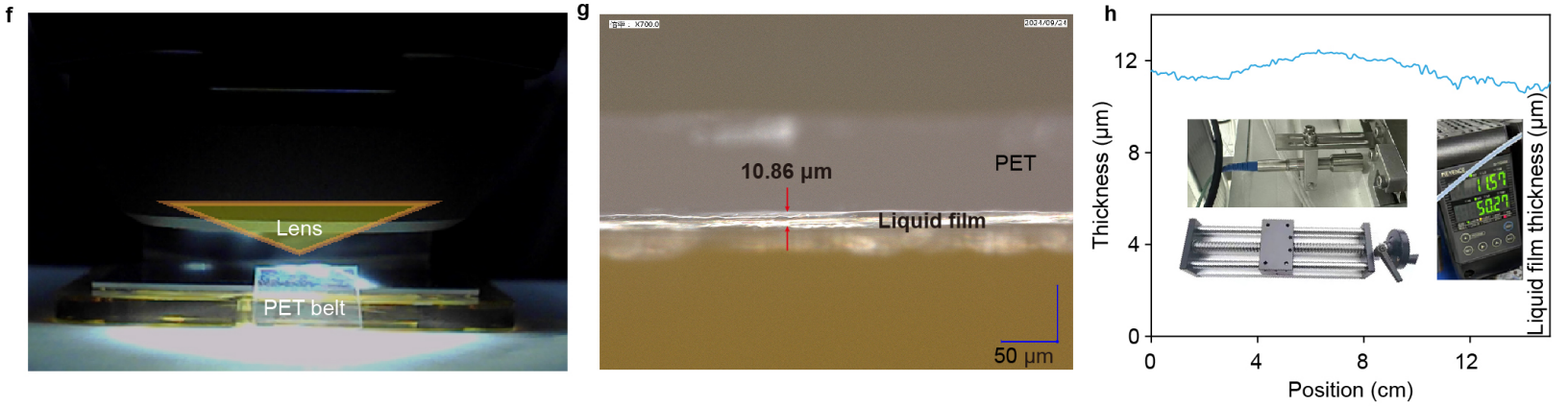


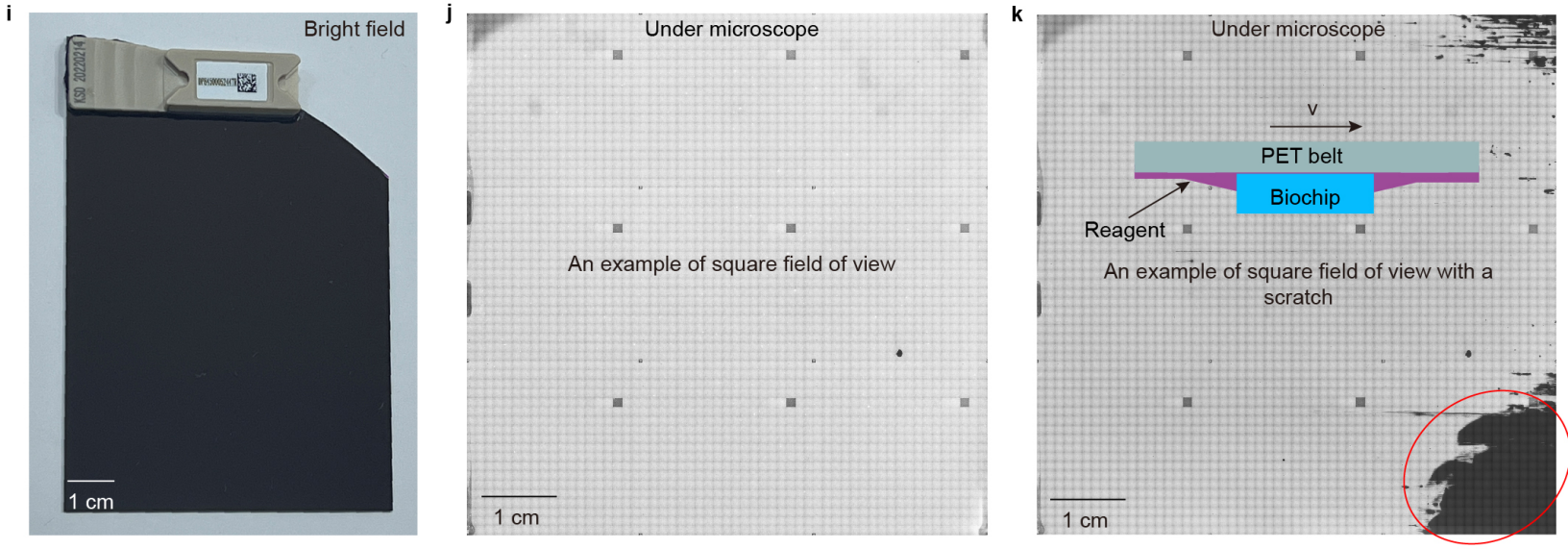
e

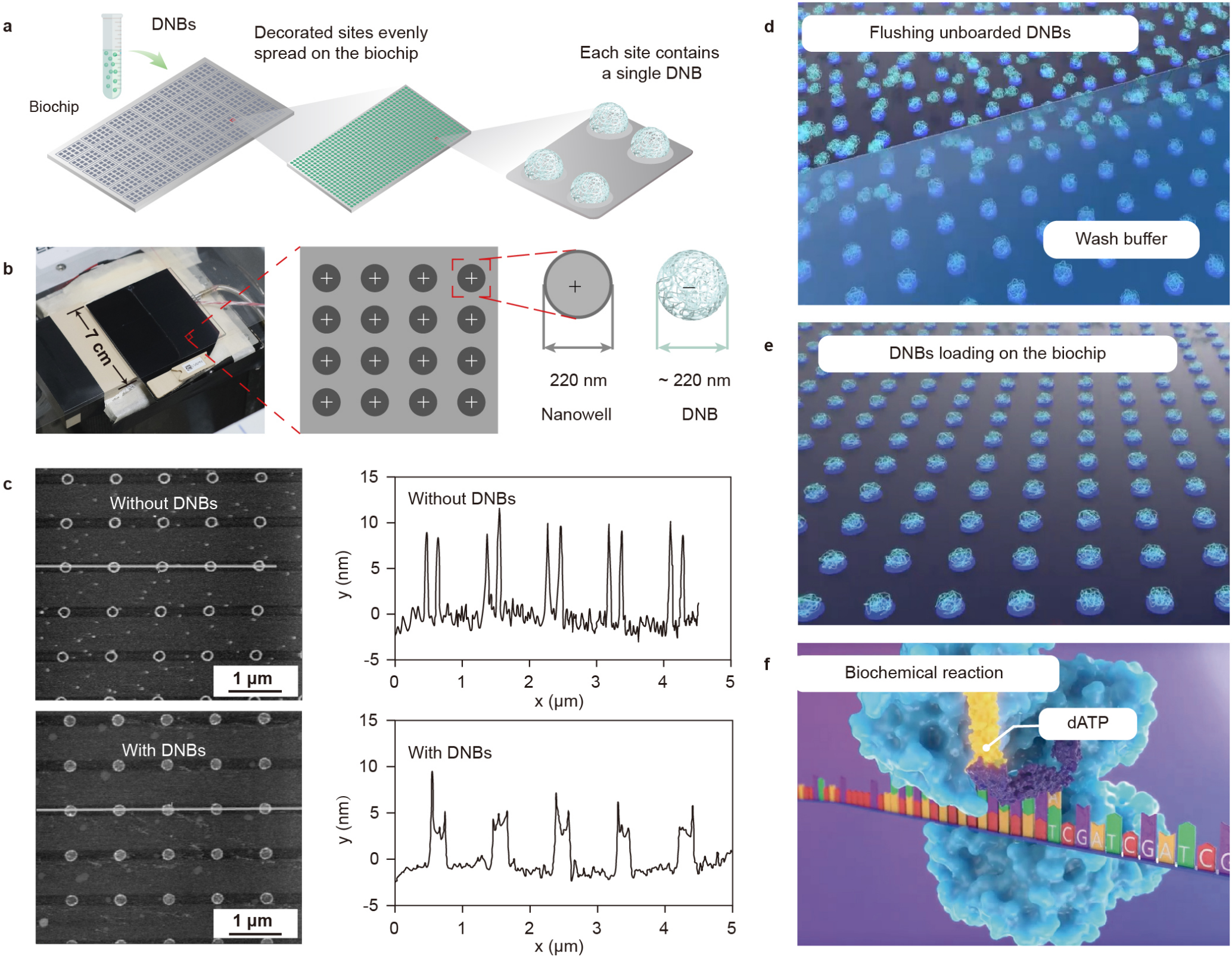


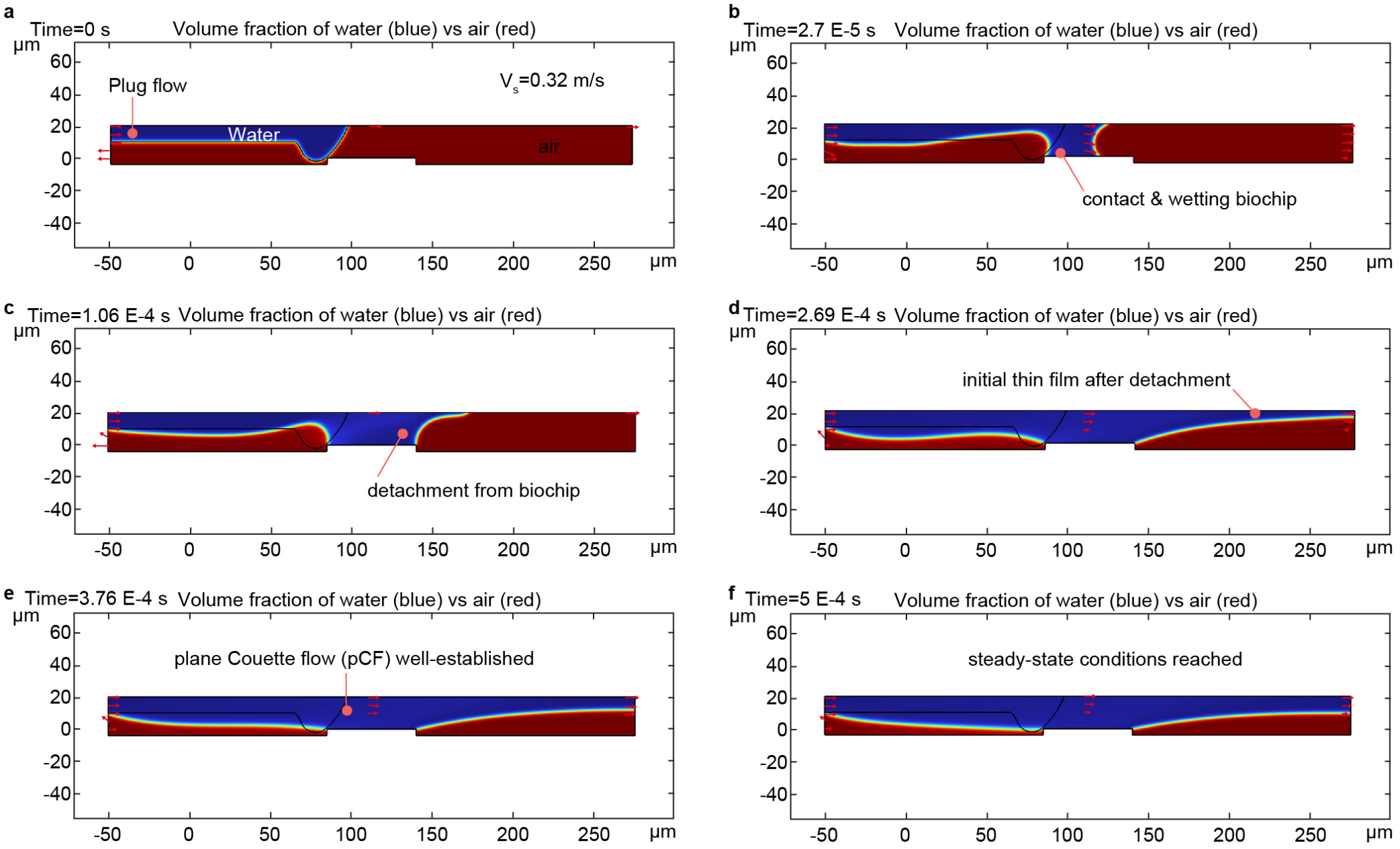

a
Aluminum Vacuum box
v
PET belt
Sensor
Biochip
b
Distance (μm)
Testing time (s)

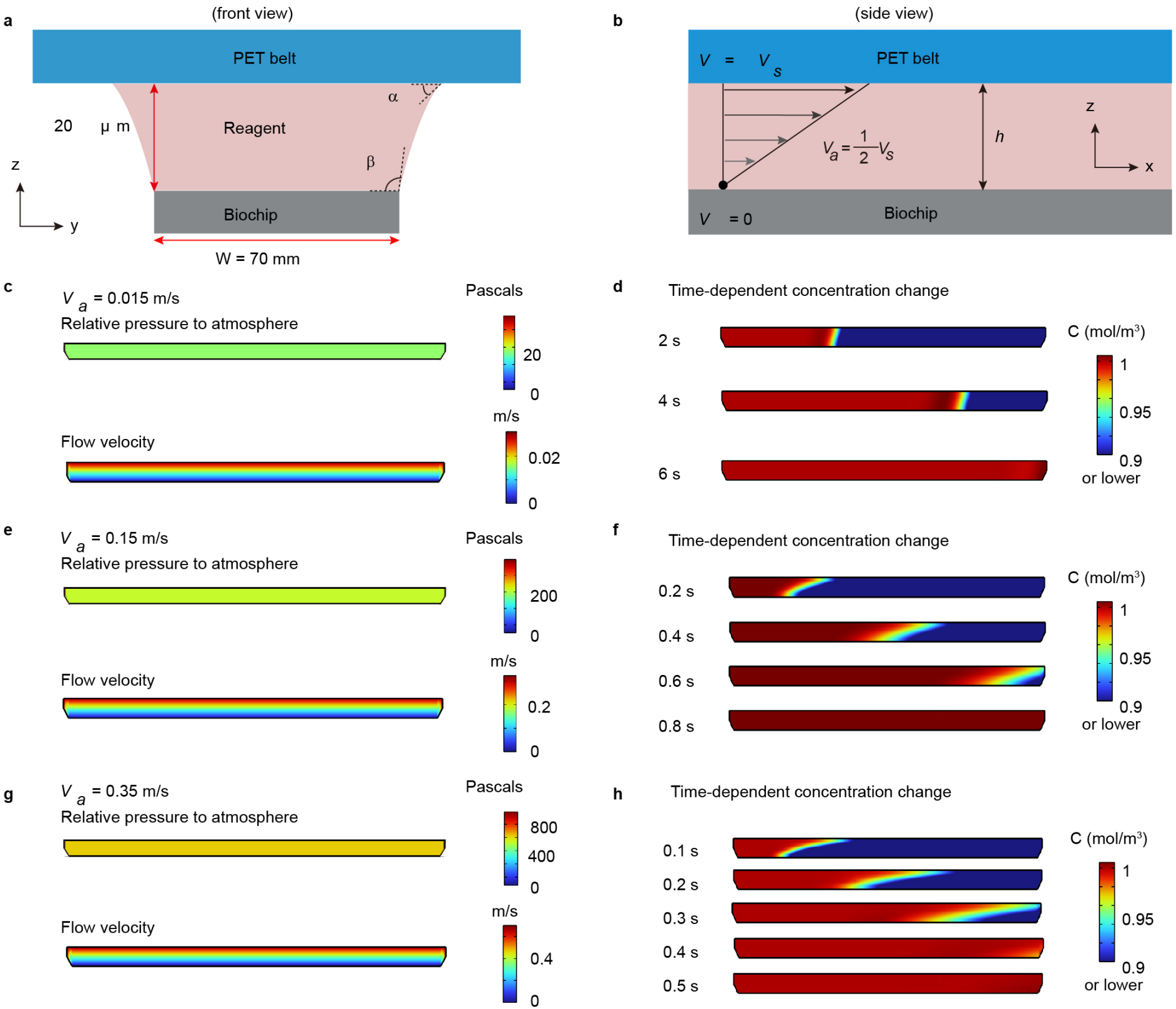

c
v
PET belt
Reagent
Gap
20 μm
Biochip
Sensor A
Sensor B
d
Thickness (μm)
Sensor A
Sensor B
Testing time (min)
e
Gap (μm)
Number of repeats


f
Lens
PET belt
g
10.86 μm
PET
Liquid film
50 μm
h
Thickness (μm)
Liquid film thickness (μm)
Position (cm)

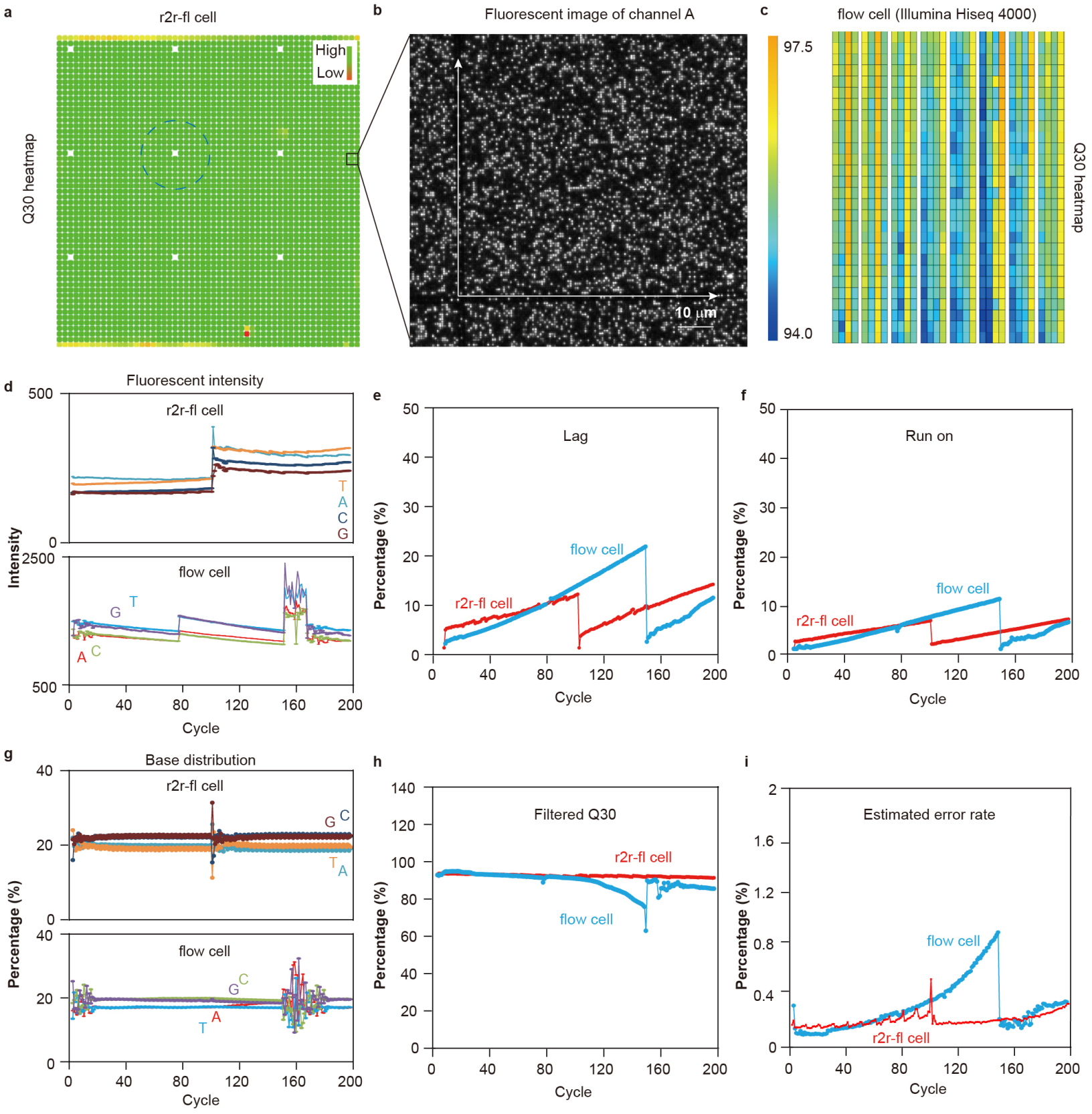

i
Bright field
1 cm
j
Under microscope
An example of square field of view
1 cm
k
Under microscope
v
PET belt
Biochip
Reagent
An example of square field of view with a scratch
1 cm

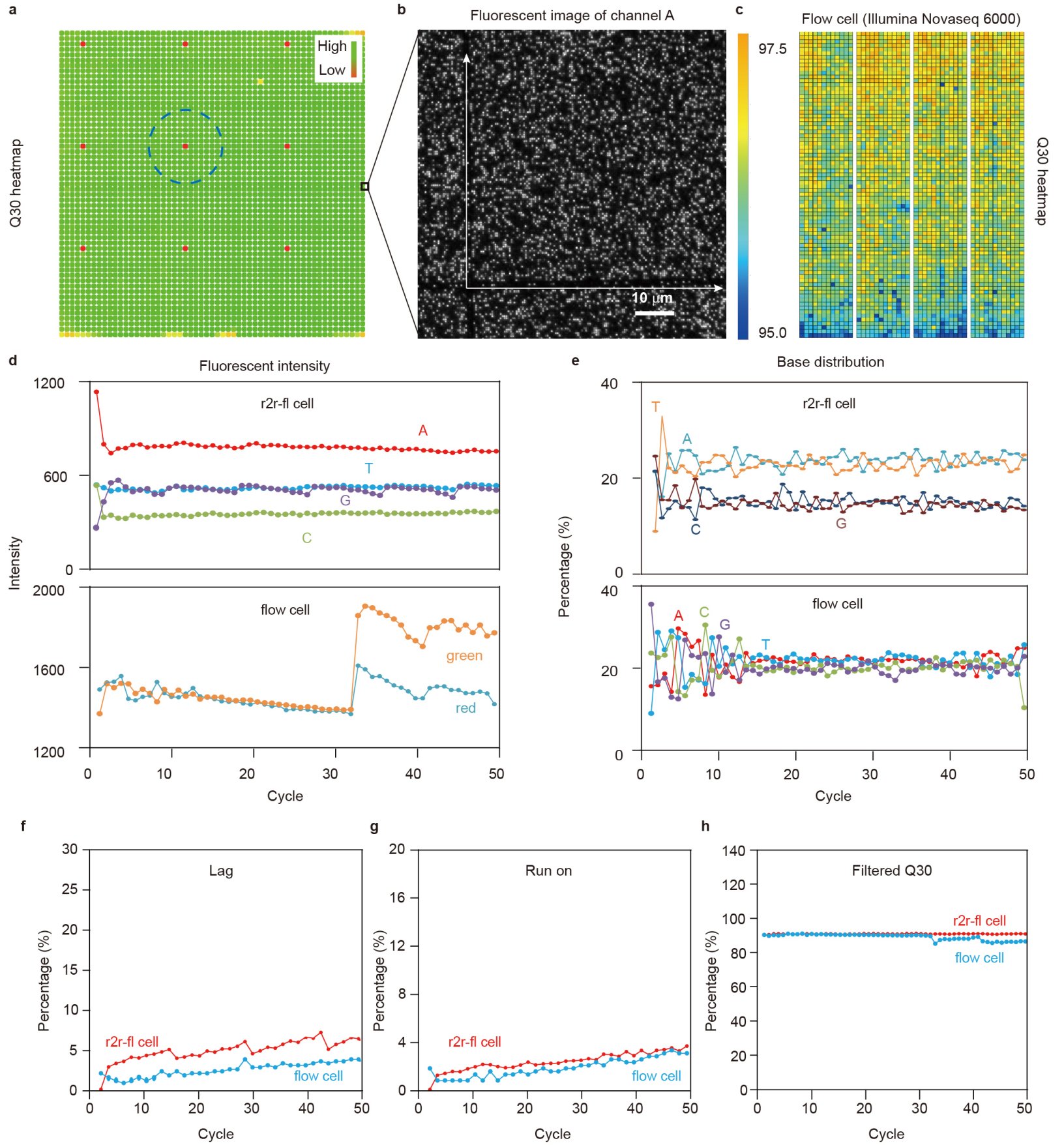

a
DNBs
Decorated sites evenly spread on the biochip
Each site contains a single DNB
Biochip
b
7 cm
220 nm
Nanowell
~ 220 nm
DNB
c
Without DNBs
1 μm
With DNBs
1 μm
Without DNBs
With DNBs
y (nm)
x (μm)
d
Flushing unboarded DNBs
Wash buffer
e
DNBs loading on the biochip
f
Biochemical reaction
dATP

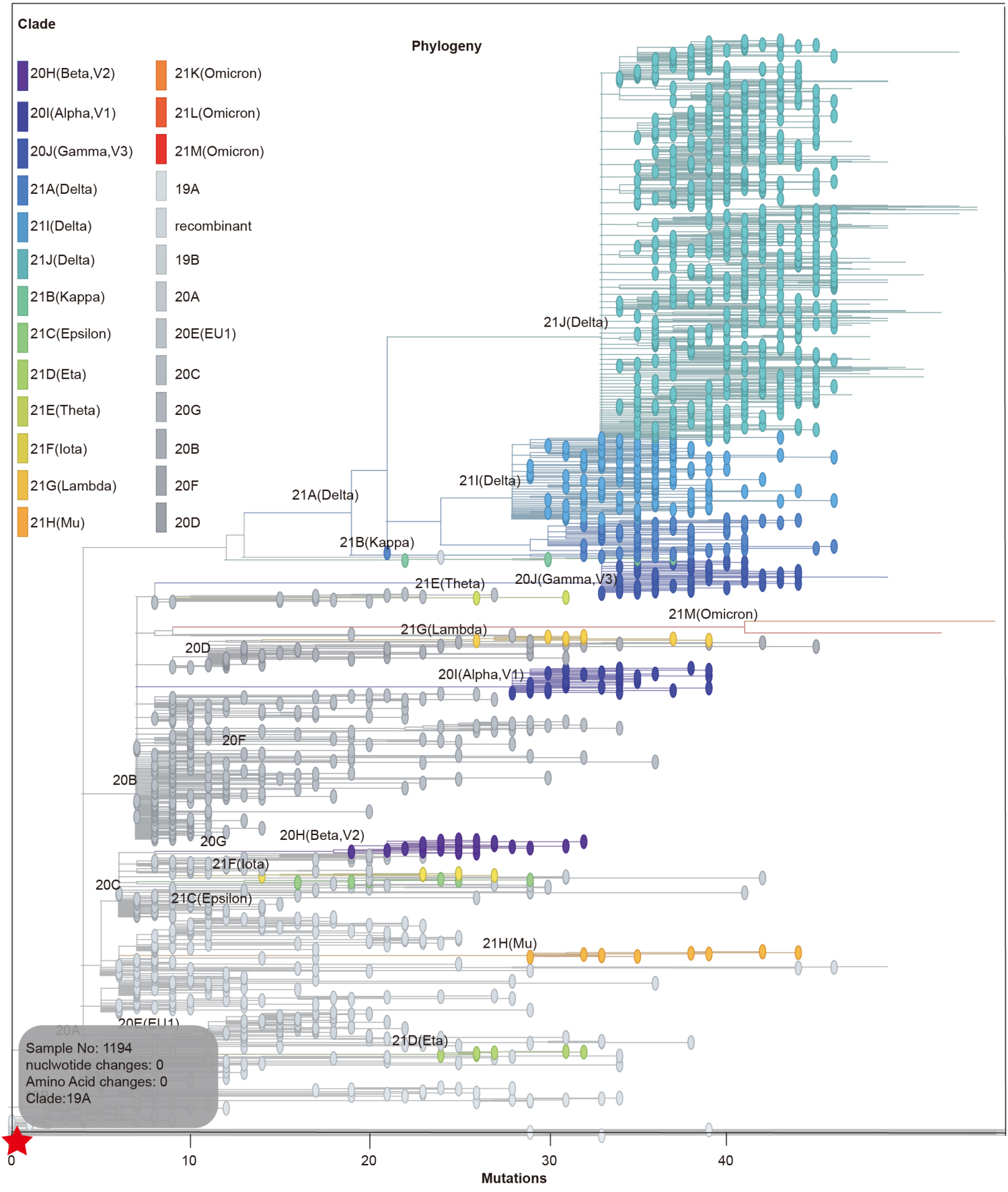

a
Time=0 s
Volume fraction of water (blue) vs air (red)
μm
Plug flow
V_s=0.32 m/s
Water
air
b
Time=2.7 E-5 s
Volume fraction of water (blue) vs air (red)
contact & wetting biochip
c
Time=1.06 E-4 s
Volume fraction of water (blue) vs air (red)
detachment from biochip
d
Time=2.69 E-4 s
Volume fraction of water (blue) vs air (red)
initial thin film after detachment
e
Time=3.76 E-4 s
Volume fraction of water (blue) vs air (red)
plane Couette flow (pCF) well-established
f
Time=5 E-4 s
Volume fraction of water (blue) vs air (red)
steady-state conditions reached
60
40
20
0
-20
-40
-50
0
50
100
150
200
250
μm


a
(front view)
PET belt
20 μ m
Reagent
α
β
Biochip
z
y
W = 70 mm
b
(side view)
V = V s
PET belt
h
z
x
V a = 1/2 V s
V = 0
Biochip
c
V a = 0.015 m/s
Relative pressure to atmosphere
Pascals
20
0
Flow velocity
m/s
0.02
0
d
Time-dependent concentration change
2 s
4 s
6 s
C (mol/m³)
1
0.95
0.9
or lower
e
V a = 0.15 m/s
Relative pressure to atmosphere
Pascals
200
0
Flow velocity
m/s
0.2
0
f
Time-dependent concentration change
0.2 s
0.4 s
0.6 s
0.8 s
C (mol/m³)
1
0.95
0.9
or lower
g
V a = 0.35 m/s
Relative pressure to atmosphere
Pascals
800
400
0
Flow velocity
m/s
0.4
0
h
Time-dependent concentration change
0.1 s
0.2 s
0.3 s
0.4 s
0.5 s
C (mol/m³)
1
0.95
0.9
or lower


EPSON
长光华大
IMA10T
BGI-Research


a
r2r-fl cell
High
Low
Q30 heatmap
b
Fluorescent image of channel A
10 μm
c
flow cell (Illumina Hiseq 4000)
97.5
94.0
Q30 heatmap
d
Fluorescent intensity
r2r-fl cell
T
A
C
G
flow cell
Intensity
Cycle
e
Lag
flow cell
r2r-fl cell
Percentage (%)
Cycle
f
Run on
flow cell
r2r-fl cell
Percentage (%)
Cycle
g
Base distribution
r2r-fl cell
flow cell
Percentage (%)
Cycle
h
Filtered Q30
r2r-fl cell
flow cell
Percentage (%)
Cycle
i
Estimated error rate
flow cell
r2r-fl cell
Percentage (%)
Cycle


a
Q30 heatmap
High
Low
b
Fluorescent image of channel A
10 μm
c
Flow cell (Illumina Novaseq 6000)
97.5
95.0
Q30 heatmap
d
Fluorescent intensity
r2r-fl cell
A
T
G
C
flow cell
green
red
Intensity
Cycle
e
Base distribution
r2r-fl cell
T
A
C
G
flow cell
A
C
G
T
Percentage (%)
Cycle
f
Lag
r2r-fl cell
flow cell
Percentage (%)
Cycle
g
Run on
r2r-fl cell
flow cell
Percentage (%)
Cycle
h
Filtered Q30
r2r-fl cell
flow cell
Percentage (%)
Cycle


Clade
20H(Beta,V2)
20I(Alpha,V1)
20J(Gamma,V3)
21A(Delta)
21I(Delta)
21J(Delta)
21B(Kappa)
21C(Epsilon)
21D(Eta)
21E(Theta)
21F(Iota)
21G(Lambda)
21H(Mu)
21K(Omicron)
21L(Omicron)
21M(Omicron)
19A
recombinant
19B
20A
20E(EU1)
20C
20G
20B
20F
20D
Phylogeny
21J(Delta)
21I(Delta)
21A(Delta)
21B(Kappa)
21E(Theta)
20J(Gamma,V3)
21M(Omicron)
21G(Lambda)
20D
20I(Alpha,V1)
20F
20B
20H(Beta,V2)
20G
21F(Iota)
20C
21C(Epsilon)
21H(Mu)
20E(EU1)
20A
21D(Eta)
Sample No: 1194
nuclwotide changes: 0
Amino Acid changes: 0
Clade:19A
0
10
20
30
40
Mutations