

High-throughput, low-cost and rapid DNA sequencing using surface-coating techniques

Yanzhe Qin^{1,3*}, Stephan Koehler², Shengming Zhao³, Ruibin Mai¹, Zhuo Liu¹, Hao Lu¹, Chengmei Xing³

The speed¹⁻³, expense¹⁻⁴ and throughput² of genomic sequencing impose limitations on its use for time-sensitive acute cases, such as rare^{4,5} or antibiotic resistant infections⁶, and large-scale testing that is necessary for containing COVID-19 outbreaks using source-tracing⁷⁻⁹. The major bottleneck for increasing the bandwidth and decreasing operating costs of next-generation sequencers (NGS) is the flow cell that supplies reagents for the biochemical processes; this subsystem has not significantly improved since 2005¹⁰⁻¹². Here we report a new method for sourcing reagents based on surface coating technology (SCT): the DNA adhered onto the biochip is directly contacted by a reagent-coated polymeric strip. Compared with flow cells the reagent layers are an order of magnitude thinner while both the reagent exchange rate and biochip area are orders of magnitude greater. These improvements drop the turn-around time from days to twelve hours and the cost for whole genome sequencing (WGS) from about \$1000 to \$15, as well as increase data production by several orders of magnitude. This makes NGS more affordable than many blood tests while rapidly providing detailed genomic information about microbial and viral pathogens^{6,13}, cancers¹⁴ and genetic disorders for targeted treatments⁶ and personalized medicine^{6,15}. This data can be pooled in population-wide databases for accelerated research and development as well providing detailed real-time data for tracking and containing outbreaks, such as the current COVID-19 pandemic.

¹MGI, BGI-Shenzhen, Shenzhen 518083, China. ²Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA. ³BGI-Shenzhen, Shenzhen 518083, China.

26 Introduction

27 High-throughput DNA sequencing is becoming ever more commonplace for public
 28 health¹⁶⁻¹⁸, and is a powerful tool for dealing with pandemics¹⁹, such as tracing the origins
 29 of SARS-CoV-2 outbreaks^{7,20} and following the virus's evolution^{6,21,22}. In China several
 30 COVID-19 outbreaks were successfully contained using source-tracing based on NGS
 31 sequencing, because their sources were viral strains that had evolved outside of China^{8,9}.
 32 But on the scale of the ever-growing current pandemic that is approaching 100 million
 33 infections and millions of deaths²³, the limited data throughput of state-of-the-art NGS
 34 (in two days a mainstream NGS platform produces 300Gb or whole genome sequences
 35 (WGS) for only 3 people²⁴), slow turn-around times (test results typically take two or
 36 more days²⁴), and high expense of current sequencing (about \$1000 per WGS or 100Gb)
 37 limit their use for tracking and containing outbreaks¹⁹.
 38 The core component of NGS is the flow cell where the bioreactions take place, and after
 39 significant improvements in 2005 the operating costs and throughput over the last five
 40 years have remained unchanged, see Fig. S1. It is a microfluidic device with a gap of
 41 ~100 μm and contains multiple lanes of attached single-stranded DNA molecules¹² that
 42 undergo a sequence of biochemical reactions and are optically imaged for base calling
 43 after each cycle²⁵. Recent attempts at improving throughput include reducing the spacing
 44 between DNA attachment points¹²; however, this reduction is curtailed by the resolution
 45 limits of optical imaging²⁶. Another approach for improving throughput and turn-around
 46 times is increasing the size of the flow cell²⁷ and boosting its flow rate. But there are
 47 limits to the flow cell's structure: larger flow cells have a propensity for non-uniform
 48 reagent flows, see Fig. S2; higher flow rates, reduced gap and the concomitant cubic
 49 increase in pressure²⁸ causes bending and delamination of the coverslip. The fastest NGS

system can generate about 3Tb (30 WGS) per day with a typical cost of about \$1000/person and turn-around time of 2 days²⁹.

Here we present a surface-coating technique (SCT) that dramatically improves data throughput and reagent usage by rapidly applying a thin reagent coating on a large biochip covered with a microarray of 5×10^{10} DNA molecules. This increases the data size and exchange rates by orders of magnitude, see Fig.1. A PET polymeric strip coated with reagents directly contacts the biochip for fast sourcing of reagents to the DNA. Bands of reagents are arranged sequentially on the strip so that step-wise lateral displacements provide reagents for each step of the sequencing-by-synthesis biochemical reactions^{25,30}, as shown in Fig. 2a. This sidesteps issues associated with flowing reagents through a thin flow cell, such as uneven reagent distribution and high pressures necessary to quickly exchange reagents. Our prototype biochip area is 225 cm² which is two orders of magnitude larger than current single-lane flow cells and can be further increased several orders of magnitude. Moreover, the reagent layer thickness is far thinner than that of flow cells, which reduces reagent waste from 99.9% to about 99% and can possibly be reduced to about 90% (see calculation section in supplementary information); this is significant because reagents currently account for 80-90 percent of operation costs³¹. The speed of coating the biochip with reagents is 80 m/min, which is over an order of magnitude faster than that in a flow cell, see Fig. 1b, and in principle can be increased to industry levels of 300 m/min³². A robotic arm transfers the biochip from the prototype SCT subsystem to the commercial imaging subsystem (DNBSEQ-T10 by MGI) for base calling between biochemical reaction cycles, as shown schematically in Fig. 2b. The data quality is comparable to that of a flow cell, as shown in Fig. 3d. For 20 cycles our unfiltered Q30

exceeds 80% and other key parameters such as the lag and run on are acceptable, as shown in Figs. 3e and f.

Results and Discussion

The main challenges for replacing the NGS flow cell with SCT were achieving uniform coatings on the PET strip, maintaining the reagent-filled gap between the PET strip and the biochip, and mitigating effects of evaporation resulting from our open-cell approach. The SCT subsystem was designed to be compatible with base calling using our imaging subsystem (DNBSEQ-T10 by MGI) with minimal modifications by patterning a DNA array similar to commercial NGS platforms, and the main challenge was adapting a robotic arm for transferring the biochip between the two subsystems.

To create uniform reagent coatings on the corona-pretreated PET strip we used slot dies on an industrial roller and matched the reagents' surface properties to that of the strip, see Fig. 2a, S3, and S4a. There are six slot dies for each of the reagents (including buffer) required for sequencing, that apply 20 μ m reagent layers to the strip with roller speeds of 80 m/min and motor response times of 10 ms. The biochip is placed downstream from the slot dies with a gap that matches the reagent layer thickness – see Fig. 2b. To maintain the layer's uniformity, we placed adhesive tape of the same thickness at the edges of the biochip, see Fig. S5. Drying is not an issue because the tape slows evaporation and the NGS reaction steps take less than a minute.

The length of the moving reagent band necessary for supplying and flushing reagents scales with that of the biochip. We find reagent bands that are twice the length of the biochip replace previous liquid and effectively recoat the biochip, see Fig. S6, video clip 1 and simulation 1 in supplementary information. Directly coating the PET strip for

contact with the biochip is far more efficient than using flow cells: the reagent layer is much thinner (20 μm , which in principle can be reduced to 2 μm ³³) and there is no tubing that needs to be flushed, see simulation 3.

After executing a bio-chemical reaction cycle the biochip is removed from the PET strip with a robotic arm and placed in our commercial DNBSEQ-T10 imaging subsystem for base calling, see video clip 2. As the biochip is exposed to air in this step the robotic arm must move quickly to prevent drying of the DNA. During imaging the biochip is placed under an objective immersed with an imaging solution. When the biochip is neither subject to biochemical reactions nor being imaged, we place it in a buffer bath to prevent evaporation.

We verified the accuracy of the genomic sequences from our SCT subsystem. Aside from the biochip's edges where the presence of the tape disrupts the flow and heat transfer, the fluorescence signal was sufficient for the remaining 92% of the biochip for base-calling, see Figs. 3a and 3b. We verified that the fluorescent channels had good separation, see Fig. 3c. Similar to our commercial NGS platform³⁴, our Q30 remains above 80% after 20 cycles, and both lag and run on are comparable, see Figs. 3d, e and f.

Finally, we compare our method with state-of-the-art NGS genomic sequencing in terms of turn-around time, bandwidth and cost, as shown in Fig. 1b. Since the reagent exchange is an order of magnitude faster than flow cells, PE100 or 212 cycles and reads can be completed within 12 hours (8.8 hours minimum with optimistic parameters for all subsystems) rather than two days. A single lane chip system such as DNBSEQ-200 produces 300Mb per day while our single chip system produces 20 Tb, which is due to much faster reagent exchange and larger

biochip area. Behind the rollers the SCT strip has a 2m x 0.4m region available for coating biochips, which is sufficient for 266 of our 225 CM² biochips that would generate 8500Tb per day. Since the data production of a single imaging subsystems is 20Tb per day, 425 such subsystems running in parallel would be required, as shown in video clip 2. SCT has a much thinner reagent layer and unlike flow chips and has no tubing which decreases reagent usage by orders of magnitude. Since the reagent cost accounts for 80-90% of current NGS platforms, this decrease in reagent usage of SCT drops the operating costs from about \$1000 to about \$15 for WGS.

Conclusions

We developed a new NGS subsystem that replaces the flow cell with SCT where the throughput is increased by two orders of magnitude and the turn-around time has decreased from several days to 12 hours without compromising data integrity. Moreover, much less reagents are wasted using SCT, thereby reducing costs by almost two orders of magnitude. We anticipate that significant improvements to our SCT prototype can be made. If adopted, our technology would rival standard diagnostics, such as blood tests and biopsies, in terms of cost and speed while providing vastly more detailed and comprehensive information in terms of the whole genome that can be collected in massive population-wide databases³⁵. This in turn will democratize personalized medicine, dramatically accelerate the pace of research and development of therapies, and provide epidemiologists with detailed data for combating outbreaks, such as those of the current COVID-19 pandemic.

Materials and Methods

Biochips were fabricated from 8-inch silicon wafers and handles were glued on for robotic handling. Deep ultraviolet (DUV) lithography was used to fabricate a 15 x 15 cm array

of circular wells with diameter of ~250 nm with center-to-center separation of 700 nm. DNA nanoballs (DNBs) were extracted from *E. coli* ATCC 8739 through fragmentation and recursive cutting with type IIS restriction enzymes, directional adapter insertion, and replication with Phi29 polymerase³⁰. Each well was loaded with one DNB using loading and post-loading procedures, see ref 30 for details. To avoid drying the DNB-loaded biochips were stored in buffer.

We purchased corona - pretreated poly (ethylene terephthalate) (PET) rolls of from Shanghai Fuzhong Limited that were 100 kilograms of transparent, 100um thick membrane with custom cut width. The roll we use is 15 cm in radius which contains more than 1000 meters membrane and supports up to 300 cycles. Our industrial roll coater has speeds that are adjustable up to 200 m/sec with response times of 10 ms. It is equipped with 6 slot dies made by Foshan edge development mold mechanical technology Inc. The strips were under 20 kg tension and the slot dies were adjusted to create 20 µm thick layers at speeds ranging from 0.3 m/min to 80 m/min.

In order to achieve uniform coatings, we matched the reagents' chemistry to that of the polymeric strips' surface (i.e. surface wettability, surface energy and surface charge). We used Acrotest Pens to measure the surface tension and found that values below 42 mN/m and for layers thicknesses below 50 um remained uniform on the pre-treated PET strip without beading. The layer thickness was measured using the Keyence sensor LKG 150, as shown in Fig. S4.

A robotic arm by Epson moved the bio-chip between the SCT subsystem and imaging subsystem and the buffer bath for storage, as shown in Fig. 2b. Strips of 3M tape, 20 µm thick one cm wide, were stuck to the biochip and served as spacers, see Fig. S5. The industrial roller advanced the PET strip in a controlled manner to transfer the desired reagents to the DNBs attached to the biochip, which took 0.25 sec. for a roller speed of 80m/min.

To build the complementary DNA molecule for base calling we performed biochemical cycles similar to those for NGS flow cells. We were able to shorten the buffer exchange time from 1 minute to within 5 seconds for all steps, and thus one complete biochemical cycle took 1.5 minutes (including transfer time for robot).

After each biochemical cycle, the chip was transferred to the imaging subsystem of our DNBSEQ-T10 sequencer by the robot arm. Because of larger biochip, it was necessary to completely imaging all the DNBs which took less than one minute - see the video clip in supplementary information. The software from our DNBSEQ-T10 was used with no modification for base calling and other standard measurements such as lag and run on.

Conflicts of interest

The authors are employee of BGI-Shenzhen or MGIttech.co and bound by confidentiality agreements.

Acknowledgements

The work is supported by BGI-Shenzhen and MGIttech.co., China Postdoctoral Science Foundation (2019M650215) and China National Gene Bank. We thank Chutian Xing for discussion and suggestion in engineering issues, Meihua Gong for instruction in biochemical process, Wangsheng Li for his help in biochip loading, Hong Xu for help in thickness determination, Ming Ni, Dong Wei for resource providing as well as Jian Liu for support of this project.

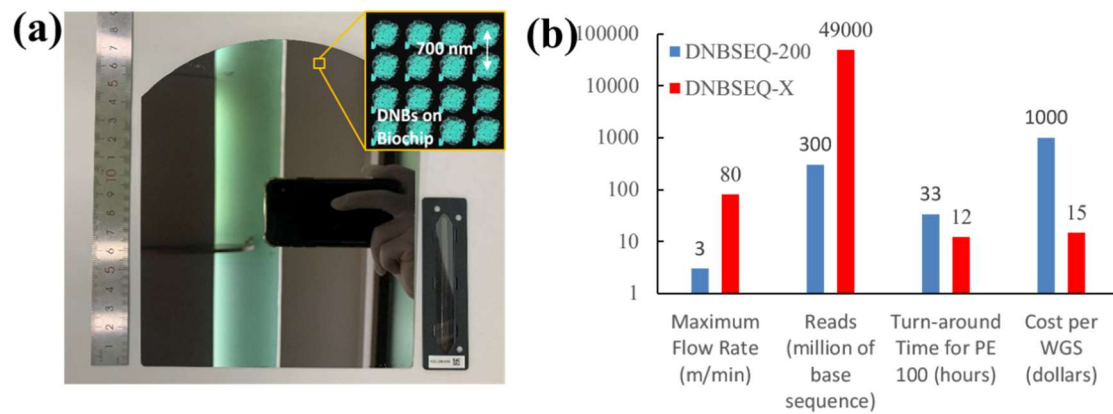


Fig. 1 Advantages of SCT. (a) Size comparison of biochips for SCT and an NGS flow cell. Inset show pattern array loading with DNBs. (b) Comparison of the flow rate, reads, turn-around time and cost for a WGS for a conventional DNBseq-200 single-lane flow cell and out SCT prototype.

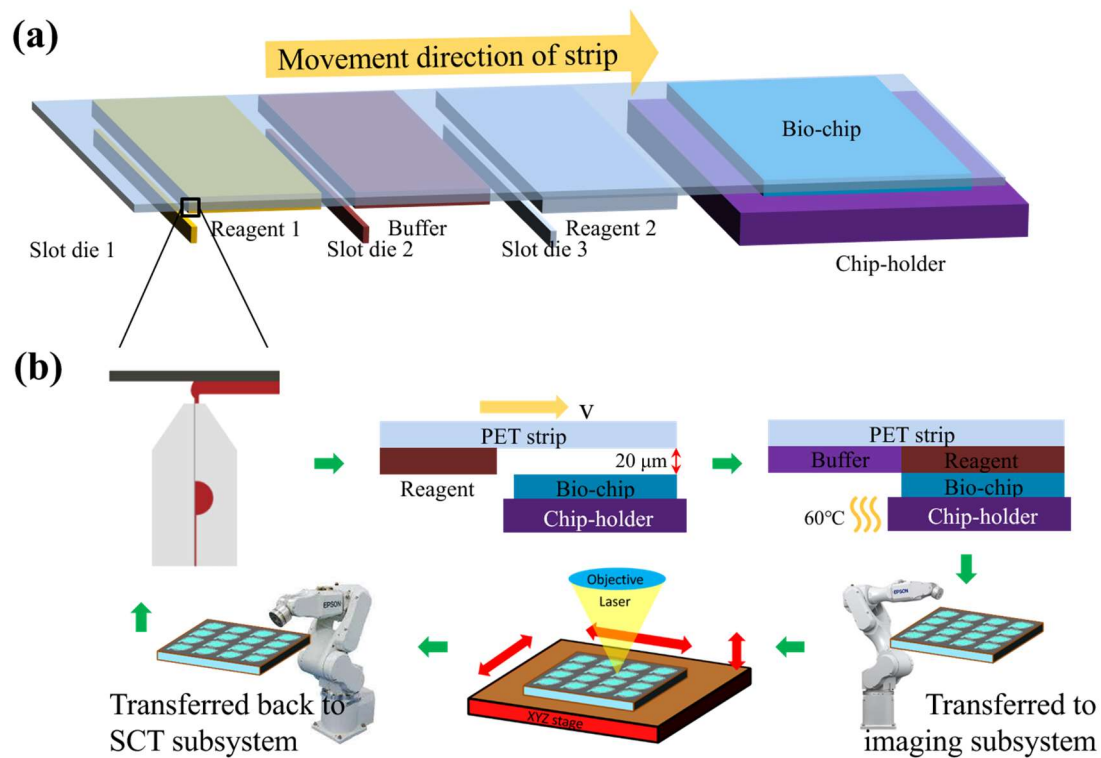


Fig. 2 Principle and set-up of the strip-coating technique (SCT) subsystem. (a) Schematic of the coating and replacement process for sequencing reagents, using a moving hydrophilic PET strip to drive flow. (b) Schematic of the workflow for SCT NGS: coating the strip, advancing the strip to source reagents and apply heat, transfer the biochip to the imaging subsystem, and returning the biochip for supplying the next reagent.

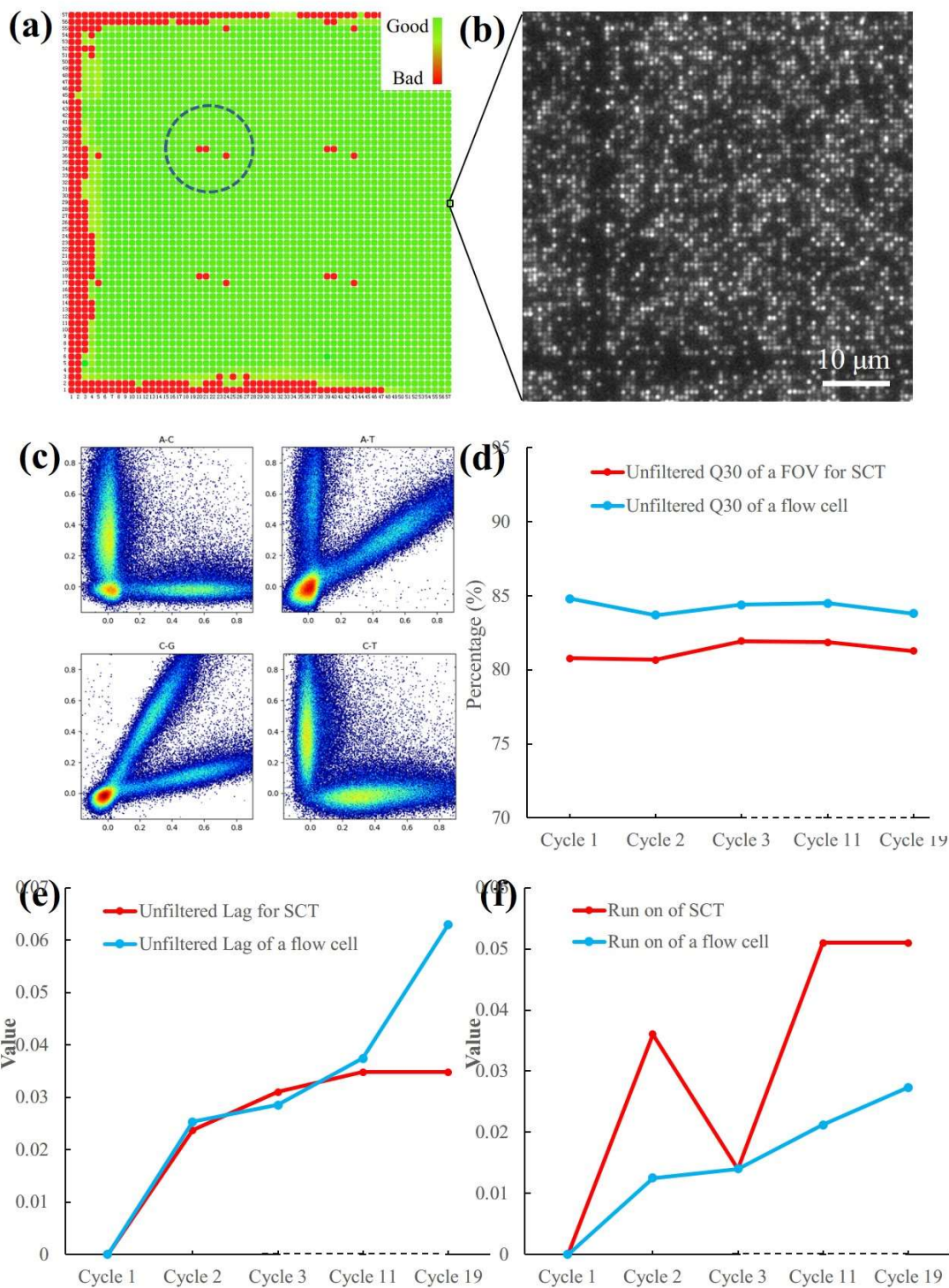


Fig. 3 Data quality achieved with our SCT prototype. (a) Base-calling information content (BIC) heatmap for a sequencing chip. Green indicates acceptable quality for base calling. The dashed white circle indicates one of the four regions used for image registration. (b) A fluorescent image zoomed in to show a square region of micropatterned DNBs (c) Crosstalk plots for the A-C, A-T, C-G, and C-T fluorescent channels, respectively, which show good channel separation. (d) Comparison of unfiltered Q30 of a FOV and that of a flow cell for the first 19 cycles. (e) The lag and (f) run on of the first 19 cycles.

200 Reference

- 201
- 202 1 Lefterova, M. I., Suarez, C. J., Banaei, N. & Pinsky, B. A. Next-Generation Sequencing for Infectious Disease
- 203 Diagnosis and Management: A Report of the Association for Molecular Pathology. *J Mol Diagn* **17**, 623-634,
- 204 doi:10.1016/j.jmoldx.2015.07.004 (2015).
- 205 2 Chiu, C. Y. & Miller, S. A. Clinical metagenomics. *Nat. Rev. Genet.* **20**, 341-355, doi:10.1038/s41576-019-
- 206 0113-7 (2019).
- 207 3 Goldberg, B., Sichtig, H., Geyer, C., Ledebor, N. & Weinstock, G. M. Making the Leap from Research
- 208 Laboratory to Clinic: Challenges and Opportunities for Next-Generation Sequencing in Infectious Disease
- 209 Diagnostics. *mBio* **6**, e01888-01815, doi:10.1128/mBio.01888-15 (2015).
- 210 4 Chai, J. H. *et al.* Cost-benefit analysis of introducing next-generation sequencing (metagenomic) pathogen
- 211 testing in the setting of pyrexia of unknown origin. *PLoS One* **13**, e0194648,
- 212 doi:10.1371/journal.pone.0194648 (2018).
- 213 5 Gu, W. *et al.* Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids.
- 214 *Nat. Med.*, 1-10 (2020).
- 215 6 Deurenberg, R. H. *et al.* Application of next generation sequencing in clinical microbiology and infection
- 216 prevention. *J. Biotechnol.* **243**, 16-24, doi:10.1016/j.jbiotec.2016.12.022 (2017).
- 217 7 Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727-
- 218 733, doi:10.1056/NEJMoa2001017 (2020).
- 219 8 Group, C.-F. R. *et al.* A Case of COVID-19 — Tianjin Municipality, China, November 7, 2020. *China CDC*
- 220 *Weekly* **2**, 884-885, doi:10.46234/ccdcw2020.241 (2020).
- 221 9 Xu, J. *et al.* A Reemergent Case of COVID-19 — Harbin City, Heilongjiang Province, China, April 9, 2020.
- 222 *China CDC Weekly* **2**, 460-462, doi:10.46234/ccdcw2020.127 (2020).
- 223 10 Shendure, J. *et al.* DNA sequencing at 40: past, present and future. *Nature* **550**, 345-353,
- 224 doi:10.1038/nature24286 (2017).
- 225 11 Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing
- 226 technologies. *Nature Reviews Genetics* **17**, 333 (2016).
- 227 12 *patterned flow cells*, <[https://www.illumina.com/science/technology/next-generation-](https://www.illumina.com/science/technology/next-generation-sequencing/sequencing-technology/patterned-flow-cells.html)
- 228 [sequencing/sequencing-technology/patterned-flow-cells.html](https://www.illumina.com/science/technology/next-generation-sequencing/sequencing-technology/patterned-flow-cells.html)> (
- 229 13 Ramachandran, P. S. & Wilson, M. R. Metagenomics for neurological infections—expanding our
- 230 imagination. *Nature Reviews Neurology* **16**, 547-556 (2020).
- 231 14 Kamps, R. *et al.* Next-generation sequencing in oncology: genetic diagnosis, risk prediction and cancer
- 232 classification. *International journal of molecular sciences* **18**, 308 (2017).
- 233 15 Rabbani, B., Nakaoka, H., Akhondzadeh, S., Tekin, M. & Mahdih, N. Next generation sequencing:
- 234 implications in personalized medicine and pharmacogenomics. *Molecular BioSystems* **12**, 1818-1830 (2016).
- 235 16 Green, E. D. *et al.* Strategic vision for improving human health at The Forefront of Genomics. *Nature* **586**,
- 236 683-692, doi:10.1038/s41586-020-2817-4 (2020).
- 237 17 Dudas, G. *et al.* Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**, 309-
- 238 315 (2017).
- 239 18 Stevens, E. L. *et al.* The Public Health Impact of a Publically Available, Environmental Database of Microbial
- 240 Genomes. *Front Microbiol* **8**, 808, doi:10.3389/fmicb.2017.00808 (2017).
- 241 19 Watson, C. How countries are using genomics to help avoid a second coronavirus wave. *Nature* **582**, 19,
- 242 doi:doi.org/10.1038/d41586-020-01573-5 (2020).
- 243 20 Zhang, Y.-Z. & Holmes, E. C. A genomic perspective on the origin and emergence of SARS-CoV-2. *Cell*
- 244 (2020).
- 245 21 Zhou, K. *et al.* Use of whole-genome sequencing to trace, control and characterize the regional expansion of
- 246 extended-spectrum β -lactamase producing ST15 *Klebsiella pneumoniae*. *Scientific reports* **6**, 1-10 (2016).
- 247 22 Tang, X. *et al.* On the origin and continuing evolution of SARS-CoV-2. *National Science Review* (2020).
- 248 23 *COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins*
- 249 *University (JHU)*, <<https://coronavirus.jhu.edu/map.html>> (
- 250 24 *nextseq 2000 specifications*, <[https://www.illumina.com/systems/sequencing-platforms/nextseq-1000-](https://www.illumina.com/systems/sequencing-platforms/nextseq-1000-2000/specifications.html)
- 251 [2000/specifications.html](https://www.illumina.com/systems/sequencing-platforms/nextseq-1000-2000/specifications.html)> (
- 252 25 Fuller, C. W. *et al.* The challenges of sequencing by synthesis. *Nat. Biotechnol.* **27**, 1013-1023 (2009).
- 253 26 Frohn, J. T., Knapp, H. F. & Stemmer, A. True optical resolution beyond the Rayleigh limit achieved by
- 254 standing wave illumination. *Proceedings of the National Academy of Sciences* **97**, 7232-7236,
- 255 doi:10.1073/pnas.130181797 (2000).
- 256 27 *T7 features*, <https://en.mgi-tech.com/Products/instruments_info/id/5> (
- 257 28 Bruus, H. *Theoretical Microfluidics*. (Oxford University Press, 2007).
- 258 29 *Novaseq specifications*, <[https://www.illumina.com/systems/sequencing-](https://www.illumina.com/systems/sequencing-platforms/novaseq/specifications.html)
- 259 [platforms/novaseq/specifications.html](https://www.illumina.com/systems/sequencing-platforms/novaseq/specifications.html)> (

- 260 30 Drmanac, R. *et al.* Human genome sequencing using unchained base reads on self-assembling DNA
261 nanoarrays. *Science* **327**, 78-81 (2010).
- 262 31 Schwarze, K. *et al.* The complete costs of genome sequencing: a microcosting study in cancer and rare
263 diseases from a single center in the United Kingdom. *Genet. Med.* **22**, 85-94, doi:10.1038/s41436-019-0618-
264 7 (2020).
- 265 32 Rosato, D. V. *Extruding Plastics: A practical processing handbook.* (1998).
- 266 33 *Slot Die Coating: How to Get Ridiculously Thin and Accurate*, <[https://www.deltamodtech.com/blog/slot-](https://www.deltamodtech.com/blog/slot-die-coating-how-to-get-ridiculously-thin-and-accurate/)
267 [die-coating-how-to-get-ridiculously-thin-and-accurate/](https://www.deltamodtech.com/blog/slot-die-coating-how-to-get-ridiculously-thin-and-accurate/)> (
- 268 34 *Sequencing Quality Scores*, <[https://www.illumina.com/science/technology/next-generation-](https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/quality-scores.html)
269 [sequencing/plan-experiments/quality-scores.html](https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/quality-scores.html)> (
- 270 35 Wilson, M. R. *et al.* Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N Engl*
271 *J Med* **380**, 2327-2340, doi:10.1056/NEJMoa1803396 (2019).
- 272 36 Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing
273 technologies. *Nature Reviews Genetics* **17**, 333-351, doi:10.1038/nrg.2016.49 (2016).

274

Electronic Supplementary Information

High throughput, rapid and low-cost DNA sequencing using surface-coating techniques

Yanzhe Qin, Stephan Koehler, Shengming Zhao, Ruibin Mai, Zhuo Liu, Hao Lu,

Chengmei Xing

*correspondence to: yanzheqin@126.com

This PDF file includes:

Figures. S1 to S6

Calculation

Supplemental figures

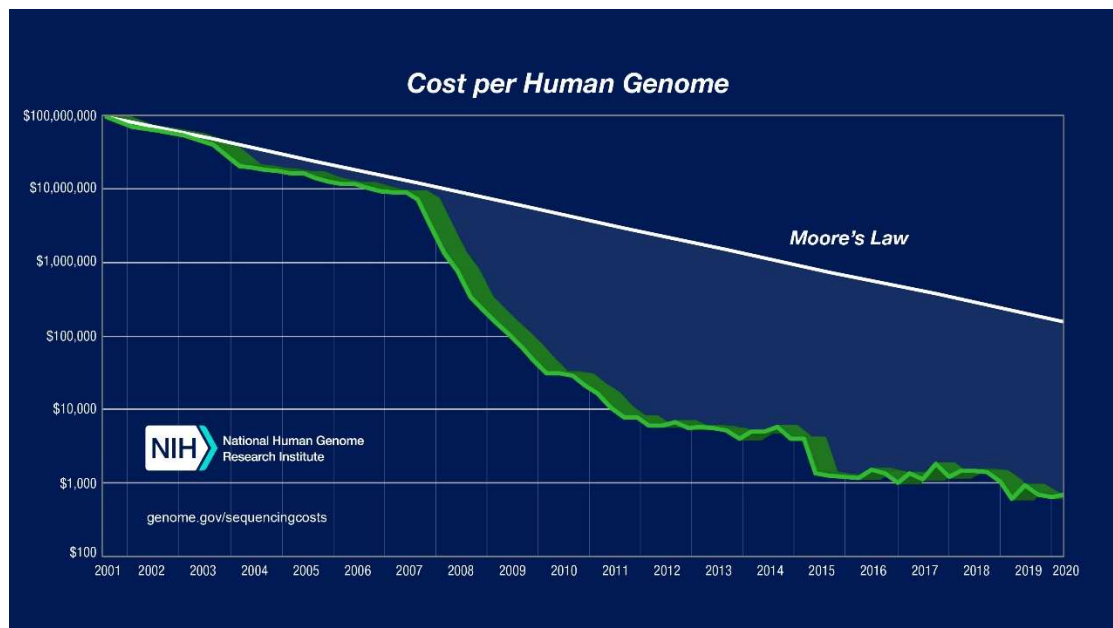


Fig. S1. The trend of cost for whole genome sequencing (WGS), and the structure of a flow cell which is the well-established method for next-generation DNA sequencing (upper right corner). The next-generation sequencing (NGS) platform first released in the mid-2000s heralded a 50,000 fold drop in the cost of human genome sequencing since the Human Genome Project³⁶. For state-of-the-art flow cells technology, more than 80% of the consumable cost arises from reagents. Courtesy: National Human Genome Research Institute, <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>. Illumina, Inc. <https://www.illumina.com/company/news-center/multimedia-images.html>.

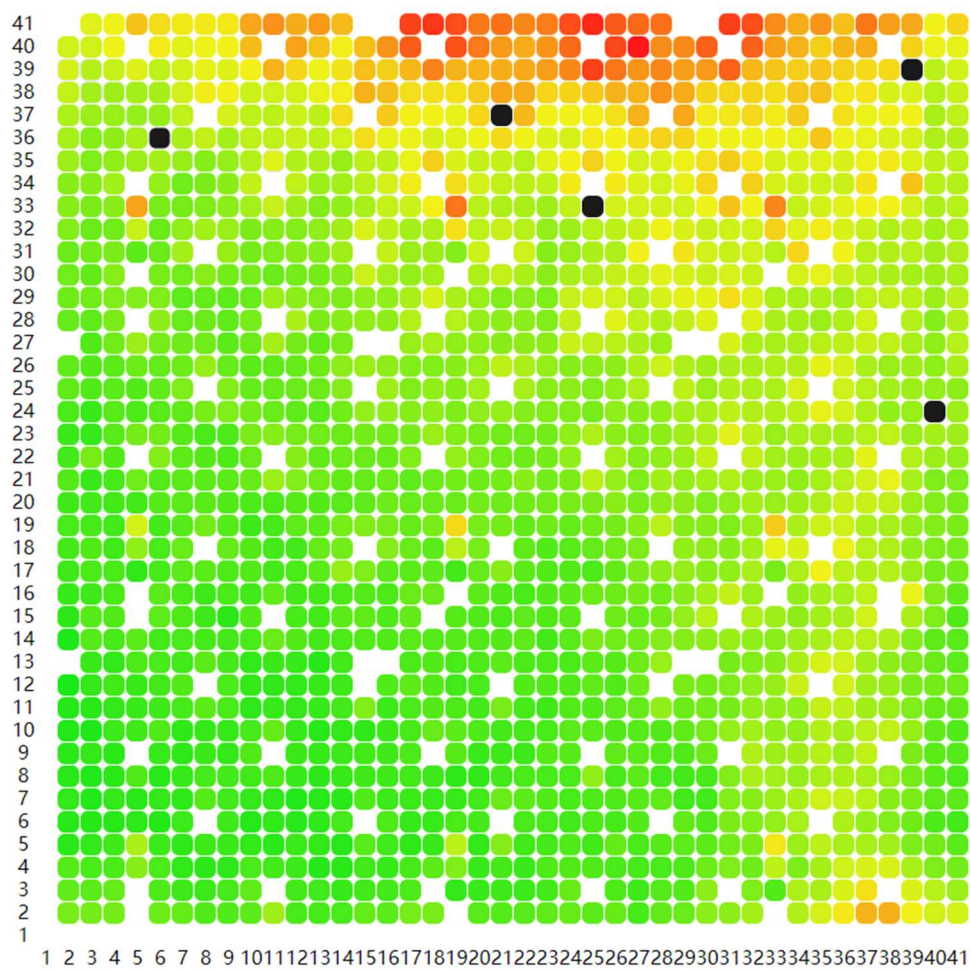


Fig. S2 A BIC (basecalling information content) heatmap of a flow cell that experienced incomplete reagent replacement. Every red dot is a FOV (field of view) that evaluated as not good enough for basecalling.

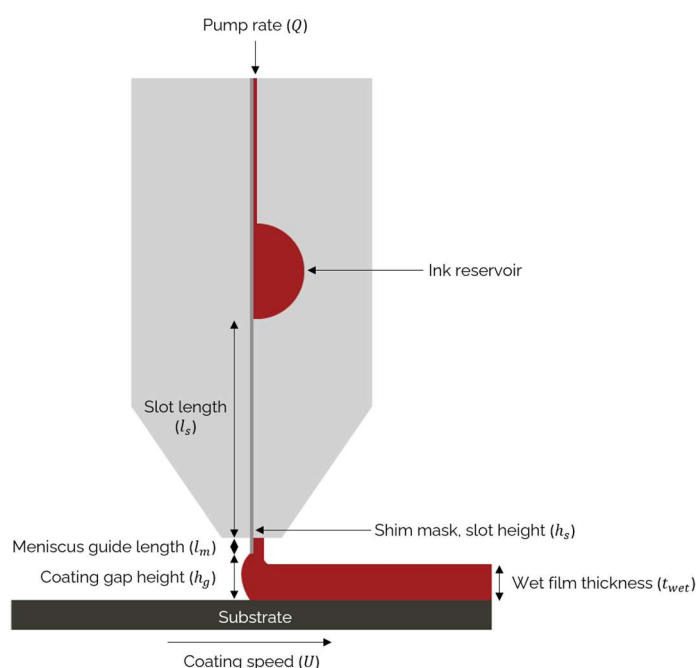


Fig. S3. The schematic of a die for thin-layer reagent coating. Courtesy: FOM Technologies.

318

PET strip with
writing of
Acrotest pen



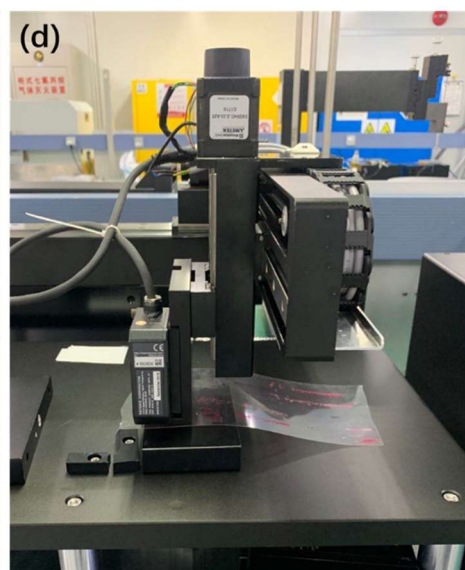
PET without
reagent coating



PET
with coating



319



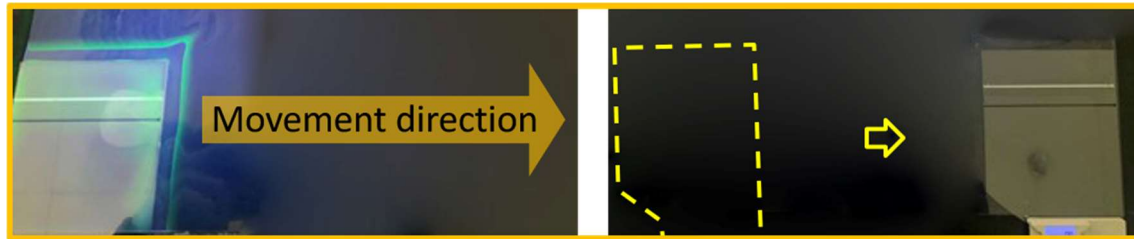
320 **Fig. S4. Uniformity and thickness test of reagent coating.** (a) PET strip
321 with writing by Acrotest pens to determine surface tension. Thickness
322 display of PET membrane (100 μ m) without (b) and with (c) reagent coating
323 using (d) Thickness determination platform based on a Keyence sensor of
324 LKG150.

325



Fig. S5. The images of a sequencing chip with two strips of 20 μ m thick tape that act as spacers.

338



339

340 **Fig. S6 Replacing a fluorescent layer on the biochip with a buffer by**
 341 **advancing the PET strip.**

342

Calculation of reagent layer thickness for one NGS cycle

Expensive contents in NGS reagents include enzymes and dNTP (N stands for A, T, C, G). Since enzymes act as catalyst in biochemical reactions, we calculate dNTP consumption as follows:

The DNA nanoballs are patterned on the biochip to form a matrix with distance of 700 nm of two adjacent DNBs, assuming we have a 7 CM by 7 CM biochip, then:

$$\text{Spot number} = 7 \text{ CM} \times 7 \text{ CM} \div (700 \text{ nm} \times 700 \text{ nm}) = 10^{10}$$

Assuming every DNB needs one fluorescent base (either A, C, T or G) in a cycle, and assuming consumption of A, C, T, G is the same for each cycle, and assuming each DNB needs ~1000 same fluorescent bases, then:

$$\text{Consumption of bases of A, C, T or G} = 10^{10} \div 4 \times 1000 = 2.5 \times 10^{12}$$

For a normal bioreaction, the working concentration of dNTP is 0.5 $\mu\text{mol/L}$, then:

$$\text{dNTP in 1 L reagent} = 6.02 \times 10^{23} \times 0.5 \times 10^{-6} = 3.01 \times 10^{17}$$

$$\text{then total bases of A, C, T, or G in 1 L reagent} = 3.01 \times 10^{17} \times 4 = 1.2 \times 10^{18}$$

For a 7 CM \times 7 CM chip,

$$\text{dNTP reagent thickness of 1 L} = 10 \text{ CM} \times 10 \text{ CM} \times 10 \text{ CM} \div 7 \text{ CM} \times 7 \text{ CM} = 20.4 \text{ CM} = 2.04 \times 10^5 \mu\text{m}$$

$$\text{So, the total number of bases in 1 } \mu\text{m thick dNTP reagent} = 1.2 \times 10^{18} \div (2.04 \times 10^5) = 5.88 \times 10^{12}$$

$$\text{Then, The required thickness of dNTP reagent} = 2.5 \times 10^{12} \div 5.88 \times 10^{12} = 0.425 \times 10^{-3} = 0.425 \mu\text{m}$$

The state-of-the-art flow-cell technique has a gap of 100 μm , assuming the buffer exchange rate is 5X, the actual reagent thickness of dNTP reagent = 100 $\mu\text{m} \times 5 = 500 \mu\text{m}$.

$$\text{The usage of reagent in state-of-the-art flow cell technology} = 0.425 \mu\text{m} \div 500 \mu\text{m} = 8.5 \times 10^{-4} = 0.08 \%$$

The reagent thickness of SCT is 20 μm and exchange ratio is 2X, then the usage of reagent is 0.425 $\mu\text{m} \div 20 \mu\text{m} \times 2 = 1 \%$

Assuming 2 μm SCT is viable, then the usage of reagent is 0.425 $\mu\text{m} \div 2 \mu\text{m} \times 2 = 10 \%$