

Open Access

Fragmentation of Surface Adsorbed and Aligned DNA Molecules using Soft Lithography for Next-Generation Sequencing

NaHyun Cho^{1,2*}, Sara Goodwin³, Julia Budassi¹, Ke Zhu¹, William Richard McCombie³ and Jonathan Sokolov¹

¹Department of Materials Science and Engineering, Stony Brook University, Stony Brook, New York 11794, USA ²Next Growth Engine R&D Coordination Division, Korea Institute of Science and Technology Evaluation and Planning, Seocho-Gu, Seoul 06775, The Republic of Korea ³Cold Spring Harbor Laboratory, 1 Bungtown Rd, Cold Spring Harbor, New York 11724, USA

Abstract

In this study, the enzymatic *in situ* cutting of linearized DNA molecules at approximately 11 kbp intervals is demonstrated using a soft lithography technique. The ultimate goal is to provide a general ordered cutting method to greatly simplify the assembly process. DNA was stretched onto PMMA (Poly methyl methacrylate) coated silicon by withdrawing the substrate from a DNA solution (a process termed "combing"). The stretched lambda DNA could be linearly cut with a soft lithography stamp used to selectively apply DNase I. After cutting the DNA on the substrate, the DNA fragments are removed from the surface by incubating PMMA in the commercial NEBuffer 3.1 at 75°C. The recovered fragments desorbed into the buffer and were sequenced using the PacBio RS II sequencer without an amplification step. The mean coverage was 2870X for the approximately 11 kbp fragmented sample and 100% of the lambda genome was sequenced. Methods to extend of the technique to ordered fragmentation are discussed.

Keywords: *In-situ* DNA fragmentation; Long-read DNA sequencing; Next-Generating sequencing; Soft lithography technology; DNA combing

Introduction

Next-generation sequencing (NGS) technology platforms have been in existence for about a decade [1,2]. The ultimate goal of NGS technology platforms is to analyze genomes with high accuracy, quickly, and at a reasonable cost. The human genome has more than 3.2 billion base pairs with many complex and repetitive regions making de novo genome assembly using current short-read technologies difficult. Pacific Bioscience's RS is one of the major recent NGS sequencing platforms, released in 2011 [1,2]. The major advantage of Pacific Bioscience's RS is the ability to sequence long contiguous DNA fragments up to 40 kbp with average reads lengths of 10-15 kbp. Comparably, Illumina platforms [1,3] (currently the most widely used short read sequencing instruments) are limited to reads of only 200 to 300 base pairs. Conventional NGS library preparation starts with fragmenting DNA into small pieces. In order to generate sequenceadaptable fragments for NGS platforms, there are two common methods: enzymatic digestion which cleaves either at specific sequence recognition sites or at random sites; or mechanical breakage with shearing forces. Two of the common methods to apply mechanical shear to the DNA phosphate backbone are sonication and needle shearing. Alternatively, Covaris (MA, USA) manufactures tubes using an adaptive focused acoustics (AFA) process that generates shearing forces through centrifugation.

Depending on which NGS platform is used, DNA is cut into fragments that range from a few hundred to a few thousand base pairs [4]. After fragmentation, DNA pieces mix randomly in the DNA buffer solution and lose all information regarding spatial organization. This loss of spatial information requires an intensive computational process to reassemble the original DNA sequence. While longer, more contiguous reads do preserve more spatial relatedness and thus improve DNA assembly [5], there is still significant room for improvement, particularly in long repetitive regions.

As Eisenstein [6] has pointed out, the current NGS sequencing technologies are based on large-scale computational reassembly of short-reads. In addition, the multiple repetitive DNA sequences make software-driven assembly of unique contigs difficult or impossible. Thus, the ability to maintain the positional order of relatively long DNA fragments from the original DNA template can provide a critical improvement for NGS platforms.

Different approaches have been developed to overcome the challenges of assembling long contigs from short-reads. For example, the 10x Genomics technology [7] (called 'linked-Reads') adds 750,000 barcodes to DNA fragments to help align sequence data to a reference genome. Another method [8] uses contiguity-preserving transposition with Tn5 and combinational indexing to perform haplotype-resolved whole-genome sequencing.

A previous approach to ordered cutting of electrostatically stretched DNA on a surface using atomic force microscopy has been reported [9,10]. Washizu's group showed that DNA molecules on a solid surface can be mechanically cut by an AFM (Atomic Force Microscope) tip and the fragments can be amplified by PCR (Polymerase Chain Reaction) with primers designed to be near cuts made along DNA of known sequence. The method described in this paper using soft lithography is simpler, faster and is able to process a larger number of DNAs in parallel. PDMS (Poly dimethylsiloxane) is a commonly used material for micro-contact printing of bio-materials [11-13], and is the material of choice in our experiments.

The aim of this paper is to demonstrate DNA fragmentation

*Corresponding author: NaHyun Cho, Next Growth Engine R&D Coordination Division, Korea Institute of Science and Technology Evaluation and Planning, Seocho-Gu, Seoul, 06775, the Republic of Korea, Tel: 1-631-632-8483; Fax: 1-631-632-8052; E-mail: ncho@kistep.re.kr

Received October 18, 2017; Accepted October 27, 2017; Published November 03, 2017

Citation: Cho N, Goodwin S, Budassi J, Zhu K, McCombie WR, et al. (2017) Fragmentation of Surface Adsorbed and Aligned DNA Molecules using Soft Lithography for Next-Generation Sequencing. J Biosens Bioelectron 8: 247. doi: 10.4172/2155-6210.1000247

Copyright: © 2017 Cho N, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

on hydrophobic coated polymer substrates using soft lithography, followed by removal into solution for use in NGS (Next-Generation Sequencing) platforms. We linearly stretch DNA molecules onto a PMMA (Poly methyl methacrylate) thin film polymer substrate spuncast onto a silicon wafer. We apply an endonuclease enzyme (DNaseI) onto the micro-patterned PDMS stamps for transfer by contact to the stretched molecules (Figure 1). After stamping with DNaseI, DNA fragments were removed by incubating PMMA into a buffer (NEBuffer 3.1, New England Biolabs). The recovered DNA fragments were treated with a FFPE (formalin-fixed, paraffin-embedded) repair kit (NEBNext* FFPE DNA Repair Mix, New England Biolabs) for damage repair. The repaired DNA was sequenced using the PacBio RS II system to confirm the quantity and quality of the DNA fragments produced using soft lithography. Fluorescence microscopy and AFM (Atomic Force Microscopy) were used to image in-situ the molecules on the substrate after cutting (Figure 1).

Material and Methods

Spin casting PMMA thin films on silicon for combing of DNA

Silicon wafers were cleaved into 2 cm by 5 cm samples and cleaned using a modified Shiraki method [14,15]. A 15 mg/mL solution of PMMA (P7490B-MMA, Polymer Source, Canada) in toluene was spun-cast onto the clean silicon wafers for 30 seconds at 2500 rpm (Headway Research PWM32 spinner). The average thickness of the PMMA thin film was found to be 700 ± 80 Å by ellipsometry (Rudolph Research AutoEL ellipsometer, USA). The PMMA-coated samples were annealed in vacuum for 4 hours at a pressure of 1×10^{-7} Torr at 130°C in a homemade ion-pumped chamber to stabilize the morphology of PMMA thin films.

Combing of fluorescent labeled DNA molecules onto PMMA

To prepare DNA solutions for the combing process, a mixture of 200 μ L of 10X DNase I reaction buffer (B0303S, New England Biolabs, USA), 200 μ L of λ DNA from NEB (500 ng/ μ L) and 7 μ L of the YOYO-1 dye (Y3601, Invitrogen, USA) stock solution were incubated for 90 minutes at 45°C to stain the DNA molecules. After fluorescent labeling, 1593 μ L of nuclease-free water was added to the DNA mixture for a final volume of 2000 μ L of 1X DNase I buffer that contains 10 mM Tris-HCl, 2.5 mM MgCl₂ and 0.5 mM CaCl₂ and a final DNA concentration of 50 ng/ μ L. To linearly comb [16,17] DNA onto a PMMA-coated substrate, a dipping apparatus was used with a computer controlled stepping motor. A teflon well was filled with 2000 μ L of the florescent DNA solution held by teflon tweezers. The computer-controlled stepping motor driving a linear stage in the dipping apparatus was used



to withdraw the substrate from the DNA solution at a rate of 1 mm per second after 30 seconds of incubation time. During the incubation time, the DNA molecules preferentially attach at the ends to the surface [16,17]. The substrate was pulled vertically at a constant speed from the solution in order to linearize the DNA molecules. After stretching the DNA molecules on the substrate, the samples were placed on a 60°C hot plate for one minute to dry any residual buffer solution.

Micron-sized patterned PDMS stamps

The micron-size pattern imprinted in photoresist-covered Si was made by UV exposure through a chromium/sodalime master (aBeam Technologies, South Korea) at the CFN (Center for Functional Nanomaterials) of Brookhaven National Laboratory, New York, USA. The Si wafers were spun-cast with positive photoresist S1805 (Dow Corning, USA), at 3000 rpm for 45 seconds, and baked at 100°C for one minute. A UV mask aligner (Karl Suss MA-6, Germany) was used to imprint micro-patterns into the photoresist. After UV exposure, the photoresist on the silicon wafer was developed in a 2:3 volumetric ratio solution of MF-312 (Rohm and Haas Electronic Materials LLC, USA) and deionized water. The photoresist patterned silicon was further etched by Reactive Ion Etching (Trion Phantom III, USA) resulting in a grating pattern with a spacing designed to produce 4.1 µm DNA fragments imprinted in the silicon substrate to a depth of approximately 0.2 µm-0.5 µm. Using a Sylgard 184 Silicone Elastomer Kit (Dow Corning, USA) the fluid polymer base and the cross-linking agent were mixed in a 10:1 weight ratio. The PDMS mix was placed in a vacuum desiccator for 1 hour to remove trapped air bubbles. Finally, the desiccated PDMS mixture was poured onto the patterned silicon wafer molds and thermally cured on a 65°C hot plate for 12 hours. After the thermal curing process, patterned PDMS stamps were cut out with a sharp knife and removed from the mold. To adhere the DNase I enzyme to the PDMS stamp, it is required to have a hydrophilic surface. The PDMS stamp as-prepared is hydrophobic. A uv-ozone cleaner (Bioforce Nanosciences, USA) was set for a 15 minute exposure to make the patterned PDMS stamp surface semi-hydrophilic. The mechanism of modifying the surface is believed to be due to the production of surface hydrophilic groups (e.g., OH) [18].

Lithographic stamping of DNase I to cut surface-bound DNA

The DNase I (M0303S, NEB, USA) cutting enzyme and 10X reaction buffer solution (B0303S, NEB, USA) were purchased from New England Biolabs. For the DNase I mixture, 140 μ L of nuclease-free water, 20 μ L of 10X DNase I reaction buffer and 40 μ L of DNase I (2000 U/mL) were mixed together. On each PDMS stamp, 25 μ L were deposited and a 10 mm by 10 mm sheet of cleanroom wipe (55% cellulose and 45% polyester, TechniCloth, Texwipe), was used to remove excessive enzyme solution on the stamp. The PDMS stamp with applied DNase I was placed on a sample of combed DNA molecules on a PMMA surface at 60°C for 15 minutes for DNase I enzymatic digestion. After the contact time, the stamp was carefully detached by tweezers and PMMA samples were heated dry at 75°C for 10 minutes to deactivate the DNase I enzyme.

AFM (Atomic Force Microscopy)

AFM images of fragmented DNA on the PMMA substrate were made with a Bruker Dimension Icon instrument using Peak Force-QNM mode at a scan rate of 0.5 Hz. The Peak Force Tapping mode can measure quantitatively modulus, adhesion, deformation and dissipation. The height image represents a topographical picture and can resolve individual DNA molecules on PMMA. The adhesion channel provides enough contrast to distinguish differences between DNA and PMMA regions on the surfaces.

Desorption of DNA fragments from PMMA substrates

An optimized protocol (see below) for detaching DNA fragments from substrate into a buffer solution, as a function of buffer type, time and temperature, was developed. A 2000 µL solution of 1X NEBuffer3.1 (B7203S, NEB, USA) was prepared and 2 samples of stretched DNA on PMMA (2cm by 5cm) were soaked into solution at 75°C for 20 minutes. To clean and concentrate the desorbed DNA, either magnetic bead clean up (A63880, Agencourt AMPure XP, Beckman Coulter, USA) or a spin column (D4010, Genomic DNA Clean & Concentrator[™]-10, Zymo Research, USA) were used.

Real-time PCR for measuring concentration of desorbed DNA

For real-time PCR, forward and reverse primers were designed using the software program Primer 3 [19] and synthesized in the Stony Brook DNA Sequencing Facility:

Forward Primer: 5' GGC AGA AAA CAG CCG CAT TA 3' Reverse Primer: 5' GGG CCG TTT TCA CGG TCA TA 3' The resulting amplicon is one unique fragment of 110 bp of λ DNA. The forward and reverse primers described above were used to amplify the concentrated DNA.A QuantiTect SYBR Green PCR Kit (204143, Qiagen, Germany) was used to detect florescence signals. In order to measure the concentration of the samples, 10 μL of 2X QuantiTect SYBR Green PCR Master Mix, 5 μL of DNA sample template, 3 μL of nuclease-free water, 1 μ L of forward primers(10 μ M), and 1 μ L of reverse primers(10 $\mu M)$ were mixed together and the 20 μL mixture was loaded into Multiplate 96-well PCR plates (MLL9651, Biorad, USA). The DNA Engine Opticon 2 (MJ Research, USA) was used to measure the concentration of DNA in the concentrated soaking solution. The Opticon 3monitor software was used to collect graphical and numerical data. As standard calibration samples, 200 pg, 40 pg, 8 pg, 4 pg, 2 pg and 1 pg of λ DNA were loaded into adjacent wells of the same PCR plate as the desorbed DNA sample. In order to calculate the mass of desorbed DNA, C, (Threshold Cycle) values of the calibrations were measured and compared to the sample containing desorbed DNA from the substrates. C, is the amplification cycle number of the intersection between the amplification curve and the threshold line.

Preparation of Pacific Bioscience libraries

The NEB Next FFPE Repair kit (M6630S, NEB, USA) was used to repair damage of the recovered DNA fragments, in each case all available DNA was used. The standard PacBio RS II library preparation protocol was conducted to sequence the DNA fragments from surfaces and all DNA material was used in library preparation and SMRTcell loading.

Results

DNA fragmentation using micro-patterned stamps

Atomic force microscopy was used to observe fragmentation of the combed DNA on PMMA surfaces using a 4.1 μ m patterned PDMS stamp. In our experiment, the PDMS patterned stamp causes DNase I to be applied at regular intervals along the DNA molecules, producing approximately 11 kbp sized fragments. We scanned DNase I stamped DNA fragments using atomic force microscopy in the adhesion mode. Adhesion mode images indicate the difference of adhesive force at different locations on the surface. In this mode, the dark brown colored

lines are DNA molecules and the light brown color regions are the PMMA substrate. In Figure 2, the cutting by DNase I can be seen to be successful, approximately 11 kbp (4.1 μ m) sized DNA fragments remaining on the PMMA surface. Figure 3 shows an optical fluorescent microscopy image of DNA fragments on the surface.



Figure 2: (a) AFM image of fragmented DNA molecules on PMMA surface inked withDNase I using $4.1 \mu m$ size patterned PDMS stamp (b) AFM image of patterned silicon mold used to produce $4.1 \mu m$ size fragments.





Citation: Cho N, Goodwin S, Budassi J, Zhu K, McCombie WR, et al. (2017) Fragmentation of Surface Adsorbed and Aligned DNA Molecules using Soft Lithography for Next-Generation Sequencing. J Biosens Bioelectron 8: 247. doi: 10.4172/2155-6210.1000247

Page 4 of 7

A sensitive capillary gel electrophoresis instrument (FEMTO pulse, Advanced Analytical Technologies, Paris, France) was used to measure the fragments size distribution. The mean fragment size was observed to be about 11 kbp with fragments ranging from ~17,000 bp to 1,503 bp (Figure 4). For this experiment, the DNA fragments were removed by phenol extraction (see below). Twenty samples of PMMA were used to produce enough materials for the gel electrophoresis (Figures 2-4).

Optimization of desorption of DNAs from PMMA substrates using NEBuffers

In order to find an effective desorption buffer solution for fragmented DNA molecules on PMMA, the NEBuffer set (1.1, 2.1, 3.1 and CutSmart) was studied. The set is designed for optimal activity of various types of restriction enzymes supplied by New England Biolabs. These buffers are not harmful to DNA molecules and also can be used for various enzymatic reactions on surface-absorbed DNA molecules.

Two samples of stretched DNA on PMMA (2 cm by 5 cm) were soaked in 2000 μ L of NEBuffer 1.1, 2.1, 3.1 and CutSmart at 37°C for 2 hours using the stock solution diluted by a factor of 2(5X) or 10(1X) for each buffer. For the dipping solution, 2000 μ L of 50 ng/ μ L λ DNA were used. After the 2 hour incubation time, 500 μ L of the total 2000 μ L of the desorption buffers containing the desorbed DNA were concentrated using a spin column (D4010, Genomic DNA Clean and Concentrator^{**}-10, Zymo Research, USA). The DNA was recovered from the column using 30 μ L of elution buffer. In order to calculate the total mass of the desorbed DNA in the original solution volume of 2000 μ L, 5 μ L out of the eluted 30 μ L was used as input for real-time PCR. The amount of desorbed DNA measured for the 5 μ L samples was scaled to infer the amount present in the original 2000 μ L sample.

As shown in Figure 5, NEBuffer 3.1 desorbed a total of 1343 ± 134 pg of DNA into the2000 µL of 1X solution and 1533 ± 153 pg for 5X solution at 37°C. Thus, the higher concentration produced a modest 14% improvement. Since the higher concentration did not show a large increase, the 1X NEBuffer 3.1 concentration was chosen for further study (Figure 5).

Temperature Effect of DNA desorption

In order to understand the temperature dependence of the DNA

desorption from PMMA surfaces using NEBuffer 3.1, four aliquots of the 2000 μ L of the buffer were prepared for testing desorption at temperatures of 22°C, 45°C, 60°C and 75°C. Two pieces of 2 cm by 5 cm PMMA with adsorbed DNA were soaked together for 1 hour at each temperature. For DNA dipping, 2000 μ L of 10 ng/ μ L λ DNA was used. Then, 400 μ L of each soaking solution was concentrated with the Zymo GDC column and the DNA was recovered into 30 μ L of elution buffer. In order to determine the total mass of the desorbedDNA at each temperature, 5 μ L out of the eluted 30 μ L was used for real-time PCR.

As the temperature increased, the mass of desorbed DNA from the PMMA increased significantly due to thermal activation. The improvement in efficiency of desorption is more than a factor of five between 37°C and 75°C. The results are shown in Figure 6.

Kinetic Study of DNA desorption on PMMA using 1X NEBuffer 3.1 at 75°C

In order to observe a kinetic desorption saturation curve, a kinetic study of DNA desorption using NEBuffer3.1 was run at 75°C. A sample of 1800 μ L of 1X buffer 3.1 was prepared and 2 pieces of 2 cm by 5 cm PMMA withdrawn from a 50 ng/ μ L λ DNA dipping solution and soaked for 1, 15, 30 and 60 minutes at 75°C. Five hundred microliters of the soaking solution were concentrated with a Zymo GDC column and 30 μ L was eluted from the column. For real-time PCR, 5 μ L out of the 30 μ L were used and based on the result (Data not shown), the amount corresponding to 1800 μ L was calculated (Figure 7). It is observed that between 20-30 minutes, the desorption rate reached saturation at approximately 2.6 ng per 1800 μ L of 1X 3.1 buffer at 75°C.

DNA concentration and purification (column vs. beads)

To be able to sequence the desorbed DNA, it is required to concentrate and purify the DNA in NEBuffer 3.1 after the desorption process. One method was to use a spin column, GDC (Genomic DNA Clean & Concentrator[™]-10), to remove DNase I and contaminants present in the buffer mixture. Another method was to use AMpure XP magnetic beads. For our experimental conditions, having a sufficient amount of desorbed DNA samples is critical for further sequencing preparation. Thus, finding the optimal method to recover DNA after the desorption process is important. In order to check the efficiency of



Page 5 of 7







Figure 6: Temperature dependence of DNA desorption from PMMA using 1X NEBuffer 3.1.



the two concentrating methods, we prepared three aliquots of 250 ng λ DNAin 25 µL. Two aliquots were concentrated with either GDC or AMpure magnetic beads, while the third was used as a calibration. The concentrated λ DNA aliquots were loaded in a 1% agarose gel(1613015, Biorad, USA) and the gel was stained with a 1:10,000 dilution of SYBR Gold dye (S11494, Life Technology, USA) in 1X TAE buffer for 30 minutes. The stained gel was imaged with the Dark Reader blue trans illuminator (DR22A, Clare Chemical Research, USA). A digital camera (Coolpix P310, Nikon, Japan) was used to image the gel. Image J was used to quantify fluorescence intensities of the samples. The concentrated sample using GDC recovered 76 ± 8 ng of the original 250 ng, so the efficiency of GDC was 30.4%. The amount of DNA recovered using the magnetic beads was 222 ± 15 ng, so the efficiency of using AMpure bead was 89%.

Sequencing results using fragmented DNA from surfaces

In order to obtain enough DNA for PacBio library preparation, twelve PMMA samples (2 cm by 5 cm) were dipped in a 50 ng / μ L λ DNA solution and fragmented with DNase I enzyme applied to the 4.1 µm fragment size PDMS stamps. After the fragmentation with the stamps, two containers of 2000 µL of 1X NEBuffer 3.1 desorption solutions were prepared for desorption of the fragmented DNA. In each container, 2 PMMA samples were soaked in 2000 µL of the buffer at 75°C for 30 minutes and for every further 30 minutes, another 2 PMMA samples were soaked in the same buffer. Desorbed DNA fragments were purified with magnetic beads (A63880, Agencourt AMPure XP, Beckman Coulter, USA). The concentration of the desorbed fragments was measured as 3.4 ng/µL in a sample volume of 120 µL as determined by a UV-Vis spectrophotometer (Nanodrop 2000, Thermo scientific, USA). The recovered DNA was treated with the NEB Next FFPE Repair kit to repair damage. Finally, the standard PacBio RS II library preparation protocol was conducted to sequence the DNA fragments from the surfaces without using an amplification step.

The generated sequencing data were compared to the reference Lambda DNA sequence (NEB3011) as a test to ensure samples were useable. The mean length of the subreads that passed filtering of the sample was 1,537 bp and the mean mapped length was 1,388 bp. The mean coverage was 2870X and the average of reference bases called was 100%. (Table 1). Thus, the generated sequence data of the desorbed DNA from surfaces covered the entire genome of lambda DNA (48,502bp).

The lower than expected subread length could be due to inefficiencies in the method; for example the stretched DNA may elongate the DNA beyond its normal double helix state, a higher than needed enzyme concentration may have digested the DNA beyond where the PDMS stamp contacted the sample, or low input into PacBio library generation may have reduced sequencing efficiency.

Job Metric	11 kbp
Number of Bases	183,058,397
Number of Reads	30,889
N50 Read Length	1,705
Mean Read Length	1,537
Mean Read Score	0.85
Mapped Reads	28,438
Mapped Read Length of Insert	1,329
Average Reference Coverage	2870.15

 Table 1: Summary of PacBio RS II DNA sequencing data using desorbed 11 kbp cut DNA fragments in NEBuffer 3.1 and concentrated with magnetic beads.

Despite the shorter subread length this validates that this method does indeed generate DNA fragments amenable to NGS approaches and that optimizations will be needed to fully capitalize on the long read capabilities of the PacBio instrument.

We also attempted this method using a 6 kb PDMS stamp and an *E. coli* sample. In this trial we were able to generate sufficient DNA to make a PacBio library and to measure the fragment size prior to library generation. The sample was run on an Agilent Bioanalyzer instrument with a High Sensitivity chip (Agilent, Santa Clara, CA). While the DNA concentration was low, the average fragment size detected was 5,565 bp (Figure 8).

We have also developed an alternative method for removing cut DNA fragments; dissolving the PMMA substrate and DNA together followed by extraction of DNA with the standard Phenol extraction. However, this method, using 12 samples of absorbed *E. coli* DNA on PMMA, only could produce a few hundreds of picogram of DNA. This is extremely difficult to sequence without amplification. Therefore, a REPLI-g single cell kit (150343, Qiagen, Germany) was used to amplify desorbed DNA fragments using Phenol extraction. Phi 29 polymerase in the kit can amplify DNA fragments without primers, facilitating amplification of the DNA fragments removed from the substrate. Sequencing results of the amplified samples showed that the amplified *E. coli* DNA was sequenced on average 3.78X with 80% of average reference bases called).

The sequencing results for the lambda DNA samples, using desorption into buffer and magnetic bead concentration, with no amplification step, were significantly better than for the *E. coli* samples for which the DNA was recovered by dissolution of the PMMA substrate and PCI extraction, followed by an amplification step. Though part of the difference is to be attributed to the much larger size of the *E. coli* genome (4.6 Mbp, or 95 times longer than lambda), a significant part is also due to the different methods of DNA recovery and to the inherent bias of the amplification step (Table 1).

Upon sequencing of the unamplifed sample only 8 Mb of sequencing data were generated with an average subread length of 1651 and maximum subread length of ~6000 bp (Table 2). As with the 11 kb



Job Metric	6 kbp
Number of Bases	8,284,152
Number of Reads	1,587
N50 Read Length	1,988
Mean Read Length	1,651
Mean Read Score	0.85
Mapped Reads	1,364
Mapped Read Length of Insert	1,386
Average Reference Coverage	1.14

 Table 2: Summary of PacBio RS II DNA sequencing data using desorbed 6kbp cut

 DNA fragments in NEBuffer 3.1 and concentrated with magnetic beads.



stamp additional optimization will be needed to enhance sequencing results. Despite the low yield, virtually all reads mapped to the *E. coli* reference generating 1.14X genome coverage (Figure 8).

Discussion

We have demonstrated a novel method of fragmenting DNA for sequencing using soft lithography. We have successfully extracted the surface-fragmented DNA and sequenced the fragments with the Pacific Bioscience RS platform. Carrying out DNA fragmentation in situ can preserve spatial orientation of DNA fragments through either iterative DNA desorption or the combined fragmentation and barcoding that can be achieved with the use of a transposase enzyme. The method is flexible in that a wide range of specific fragment sizes or fragment size profiles (which may be conveniently varied by creating, lithographically, different PDMS stamps). The technique is bias-free and has proven capable of providing material amenable to DNA sequencing without amplification. Our method could be used to preserve spatial relatedness as well as long fragment length to provide highly contiguous fragments for genome assembly allowing for superior genomic structure resolution.

Further development of this method to enable ordered removal of the fragments is in progress. One straightforward extension would be to have fragments of graded size along the length of the DNA molecules, such that the length of the fragments would permit the ordering of the fragments along the DNA molecules (Figure 9). The fragments would be separated by length using electrophoresis. Ordered removal using localized electric field-assisted desorption or a microfluidics approach may also be feasible.

We note that ends of DNA molecules may be tethered along lines on a surface [20] and can be used to align the sequences being fragmented by our soft lithography technique. Similarly, recognition sites in long DNAs could be used as tethering points.

An alternative strategy would be to use patterned PDMS stamps

Page 7 of 7

to apply transposons (1) to insert known barcodes to the ends of the linearized DNA fragments, followed by amplification and sequencing (Figure 9).

Lastly, we mention that patterns with micron-sized holes may be used to create nanoliter volume wells into which the fragments may be desorbed and which could serve as nanoreactors for various amplification schemes [21].

Acknowledgments

Lithography research was carried out at the Center for Functional Nanomaterials, Brookhaven National Laboratory, which is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-SC0012704.

We would like to thank to Robert Wappel, Panchajanya Deshpande from Cold Spring Harbor Laboratory, Dr. Jiang Chen and Dr. Yeun Ja Choi from Departments of Pathology and Dermatology at Stony Brook Medicine and undergraduate student, Donald Liu at Stony Brook University.

Funding

Support from the National Science Foundation (NSF DMR-0606387) is gratefully acknowledged.

References

- Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics 17: 333-351.
- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, et al. (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC genomics 13: 341.
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, et al. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal 6: 1621-1624.
- Metzker ML (2010) Sequencing technologies the next generation. Nature reviews. Genetics 11: 31-46.
- Roberts RJ, Carneiro MO, Schatz MC (2013) The advantages of SMRT sequencing. Genome Biology 14: 405.
- 6. Eisenstein M (2015) Startups use short-read data to expand long-read sequencing market. Nature Biotechnology 33: 433-435.
- Zheng GX, Lau BT, Schnall-Levin M, Jarosz M, Bell JM, et al. (2016) Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. Nature Biotechnology 34: 303-311.

- 8. Amini S, Pushkarev D, Christiansen L, Kostem E, Royce T, et al. (2014) Haplotype-resolved whole-genome sequencing by contiguity-preserving transposition and combinatorial indexing. Nature Genetics 46: 1343-1349.
- Kurosawa O, Washizu M (2007) Dissection, acquisition and amplification of targeted position of electrostatically stretched DNA. Journal of Electrostatics 65: 423-430.
- Li B, Zhang Y, Yan SH, Lu JH, Ye M, et al. (2007) Positioning scission of single DNA molecules with nonspecific endonuclease based on nanomanipulation. Journal of the American Chemical Society 129: 6668-6669.
- Gad M, Sugiyama S, Ohtani T (2003) Method for patterning stretched DNA molecules on mica surfaces by soft lithography. Journal of Biomolecular Structure & Dynamics 21: 387-393.
- Lange SA, Benes V, Kern DP, Horber JK, Bernard A (2004) Microcontact printing of DNA molecules. Analytical Chemistry 76: 1641-1647.
- Thibault C, Le Berre V, Casimirius S, Trevisiol E, Francois J, et al. (2005) Direct microcontact printing of oligonucleotides for biochip applications. Journal of Nanobiotechnology 3: 7.
- Okumura H, Akane T, Tsubo Y, Matsumoto S (1997) Comparison of conventional surface cleaning methods for Si molecular beam epitaxy. Journal of the Electrochemical Society 144: 3765-3768.
- Shin K, Hu X, Zheng X, Rafailovich M, Sokolov J, et al. (2001) Silicon oxide surface as a substrate of polymer thin films. Macromolecules 34: 4993-4998.
- Allemand JF, Bensimon D, Jullien L, Bensimon A, Croquette V (1997) pHdependent specific binding and combing of DNA. Biophys J 73: 2064-2070.
- 17. Bensimon A, Simon A, Chiffaudel A, Croquette V, Heslot F, et al. (1994) Alignment and sensitive detection of DNA by a moving interface. Science 265: 2096-2098.
- Efimenko K, Wallace WE, Genzer J (2002) Surface modification of Sylgard-184 poly (dimethyl siloxane) networks by ultraviolet and ultraviolet/ozone treatment. J Colloid Interface Sci 254: 306-315.
- 19. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, et al. (2012) Primer3--new capabilities and interfaces. Nucleic Acids Res 40: e115.
- Collins BE, Ye LF, Duzdevich D, Greene EC (2014) DNA curtains: novel tools for imaging protein-nucleic acid interactions at the single-molecule level. Methods Cell Biol 123: 217-234.
- Marcy Y, Ishoey T, Lasken RS, Stockwell TB, Walenz BP, et al. (2007) Nanoliter reactors improve multiple displacement amplification of genomes from single cells. PLoS Genet 3: 1702-1708.