nature nanotechnology

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NNANO.2017.176

In the format provided by the authors and unedited.

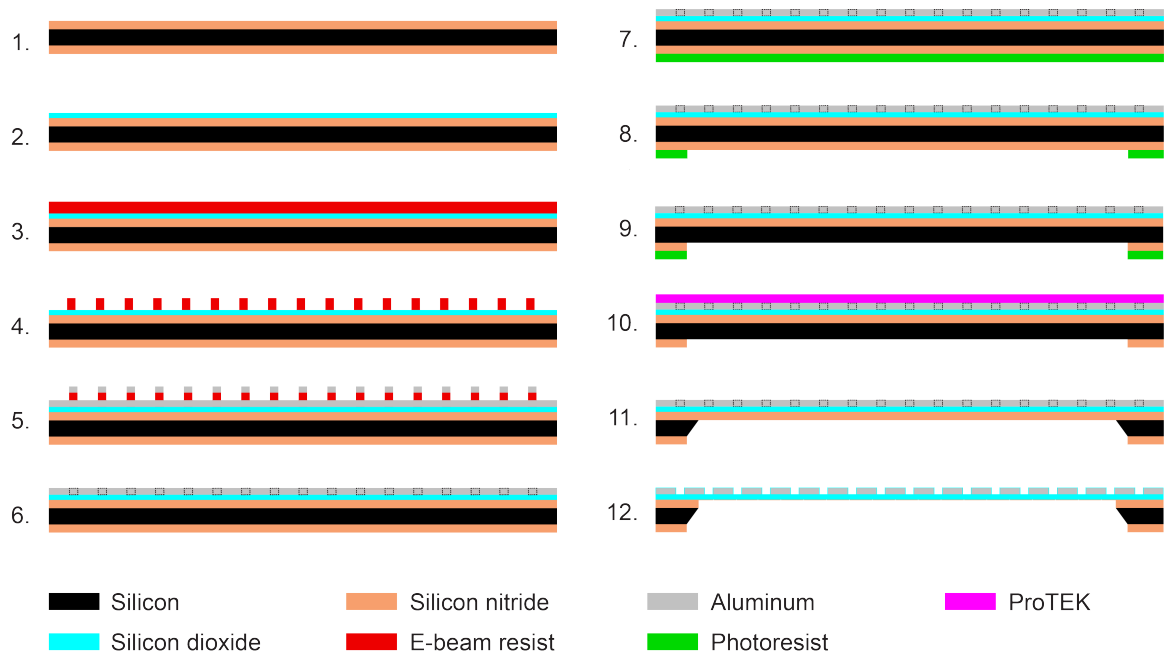
Length-independent DNA packing into nanopore zero-mode waveguides for low-input DNA sequencing

By

Joseph Larkin,[†$] Robert Y. Henley,[†$] Vivek Jadhav,[†$] Jonas Korlach,[#] and Meni Wanunu[†‡*]

Sections

1. NZMW fabrication



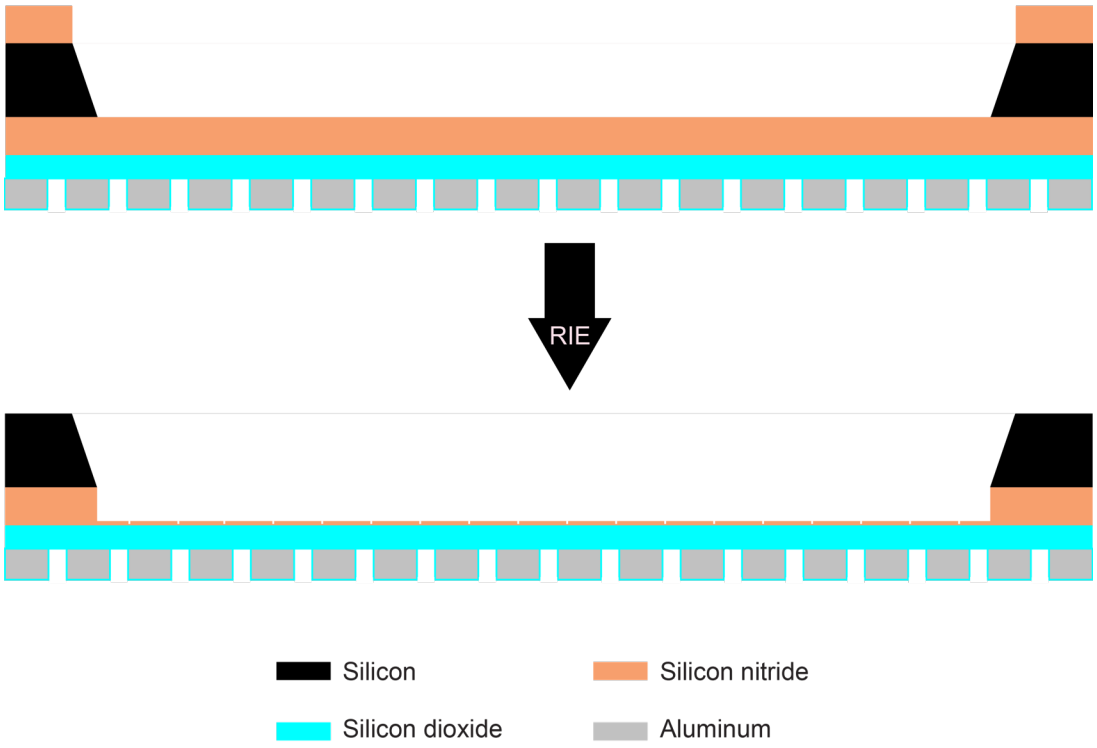
Supplementary Figure 1: Stepwise Schematic representation of NZMW fabrication process: 1. LPCVD silicon nitride coating of silicon wafer. 2. SiO_2 coating on silicon nitride on one side of wafer. 3. Spin AR-N 7520.11 negative e-beam resist. 4. Expose using EBL and develop resist. 5. Thermally evaporate 100 nm of aluminum. 6. Liftoff resist. 7. Spin S1818 positive photoresist on reverse side of the wafer. 8. Expose and develop positive photoresist. 9. Etch silicon nitride windows with reactive ion etching. 10. Spin and bake protective ProTEK on ZMW surface (on top of aluminum). 11. KOH etch silicon. 12. Remove ProTEK and deposit ALD SiO_2 .

To fabricate NZMWs we use 200 µm-thick, 4" <100> Si wafers. The wafers were cleaned with a standard process, rinsed with DIH_2O and dried with an N_2 gun. Prior to fabrication, 35 nm of low stress silicon nitride (SiN_x) was deposited on either side of the wafer using low pressure chemical vapor deposition (Fig S1, Step 1). We deposited 20 nm of SiO_2 via atomic layer deposition with Bis(diethylamino)silane (BDEAS) as a precursor and ozone as an oxidizer (Fig S1, Step 2). One side of the wafer will be used to fabricate ZMWs on SiN_x membranes, which will be etched from the reverse side with KOH.

ZMWs were then fabricated using e-beam lithography (EBL) using AR-N 7520.11 negative resist (Allresist GmbH). E-beam resist was spun onto one side of the wafer (3300 rpm for 45 sec) and baked (60 sec at 90 C) (Fig S1, Step 3). Using a JEOL 6300FS e-beam writer, a ZMW pattern containing an array of circles (80nm diameter) and an alignment mark was exposed in the resist. Alignment marks are used to ensure that the ZMW arrays line up with a backside photolithography mask for etching down to SiN_x membranes. After exposure the negative beam resist was developed with MF321 solution (90 sec), rinsed with DIH_2O (120sec) and dried with nitrogen. To ensure full removal of resist, wafers were descummed in downstream oxygen plasma (100 watts, 60 sec) (Fig S1, Step 4) and 100 nm of thermally evaporated aluminum was deposited immediately (Fig S1, Step 5). To form ZMWs, a lift-off process was carried out in sonicated 1165 Remover solution (MicroChem, 3 hrs at 68 C) (Fig S1, Step 6) and later it was kept overnight in

the same solution to ensure complete lift-off. Wafers were rinsed thoroughly with DIH_2O and descummed again with oxygen plasma. To verify proper metallization and lift-off, ZMW patterns were checked with a scanning electron microscope (SEM) (SEM image shown in figure 1a of main text) prior to photolithography on the back side of the wafer. A defined pattern was photomasked then written with photolithography on the back side each wafer for KOH etching of silicon. Positive photoresist (Shipley Microposit S1818) was spun (4500 rpm, 60 sec) on the back side of the wafer and baked (115 C, 60 sec) (Fig S1, Step 7), resist was later exposed using a contact aligner (MA/BA 6, Suss MicroTech). A photomask with alignment mark was used which aligned to the features on the ZMW pattern. After the alignment resist was exposed and developed using MF321 solution (60 sec), rinsed with DIH_2O and dried with nitrogen (Fig S1, Step 8). The exposed (back) side of the wafer was treated with SF6 plasma with reactive ion etching to remove silicon nitride and expose windows for KOH etching down to free-standing SiN_x ZMW membranes (Fig S1, Step 9). Before KOH etching, ZMWs were protected from KOH with proTEK (Brewer Science) (Fig S1, Step 10). Following KOH etching and removal of proTEK in solvents (Fig S1, Step 11), we etched the backside of the ZMW membranes in SF_6 plasma (Fig S1, Step 12) to reduce photoluminescence (also see section 2, below). We then fabricated nanopores at ZMW bases using a 200 kV transmission electron beam.

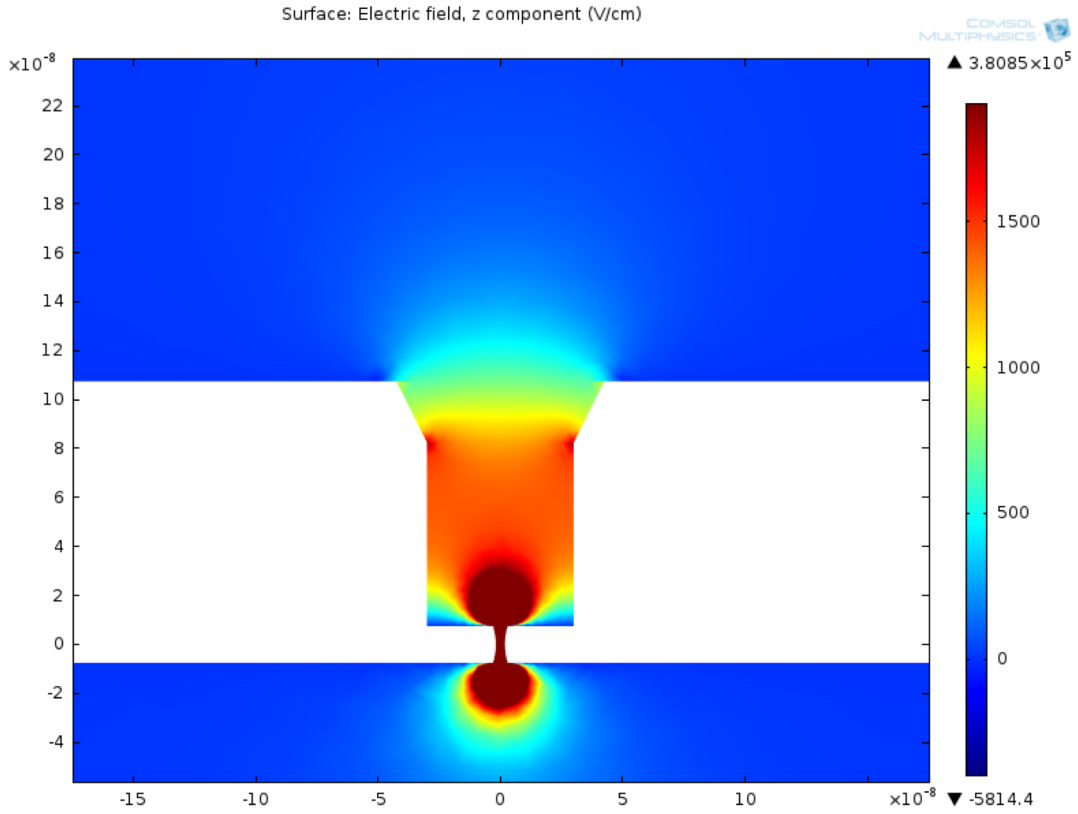
2. Reducing Nitride Photoluminescence Background



Supplementary Figure 2. NZMWs (aluminum) are first fabricated on silicon wafers with 35 nm of silicon nitride and 20 nm of atomic layer deposition silicon dioxide. After KOH etching to release membranes, most of the luminescent nitride is removed with SF_6 reactive plasma. Although the etching should suffice to remove all nitride, residual

Low-stress silicon nitride is known to exhibit photoluminescence under blue and green illumination, which can prevent single fluorophore detection[1]. In order to reduce this background, we fabricated the ZMW array on silicon dioxide. We did, however, wish to use silicon nitride as a KOH etch mask. For this reason, we deposited 20 nm of silicon dioxide on a wafer with 35 nm of low-stress silicon nitride. After fabricating ZMWs and back etching, we removed the silicon nitride with reactive SF_6 plasma in a Technics micro-RIE 800 system under the following conditions: 300 mTorr, 150 W, and etch times generally 1-2 minutes. Longer times often resulted in fragile membranes. RIE resulted in the reduced luminescence depicted in Figure 1d of the main text. The concept is illustrated schematically in Figure S2. The RIE removes much of the nitride, reducing photoluminescence, and the thin SiO_2 layer is enough to support the ZMW membrane.

3. Numerical solution for the NZMW's DC electric field

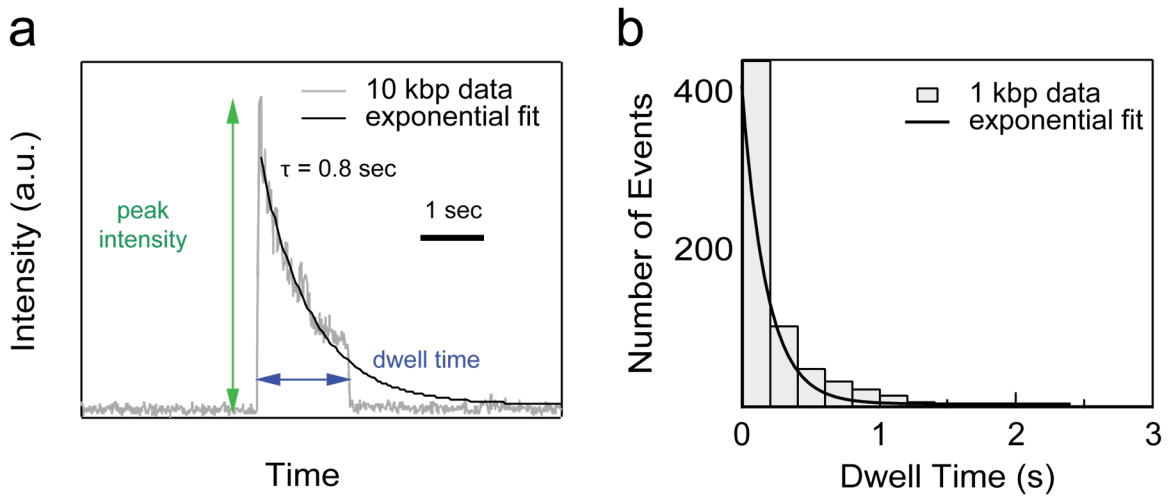


Supplementary Figure 3. The numerical solution for the DC electric field in and around a 100 nm-high ZMW with a 3 nm nanopore under 500 mV bias. To match the observed ZMW geometry, it has a 60 nm base diameter with a truncated cone structure having an 85 nm outer diameter. The electrolyte contains a 400 mM KCl concentration and the pore has a 5 nm effective height. Color scale represents the magnitude of the electric field's z component

4. Analysis of YOYO-1-stained DNA

To obtain the data presented in Figure 2 of the main text we first extracted the fluorescence traces from each ZMW in ImageJ. The peak fluorescence intensity for each YOYO-1 spike was calculated as the difference between the average baseline level and the initial height of the fluorescence spike (see Figure S4a). We then analyzed these traces with Pythion software (https://github.com/rhenley/Pyth-Ion) to extract inter-event times and dwell times for each event. For Pythion to properly analyze these traces, we had to invert the data and chose an appropriate threshold for the beginning and end of a molecular event. This resulted in an underestimate of

the dwell time of any molecules that photobleached before escaping the NZMWs. Figure S4a illustrates the dwell-time for an example 10 kbp spike. Additionally, we fit the exponential decay for these fluorescence spikes as shown in Figure S4a. The mean and standard deviation of these decay times are reported in the main text. After collecting dwell times, we assigned a characteristic time scale for each DNA length and applied voltage by fitting an exponential to the dwell time histograms, as depicted in Figure S4b for 1 kbp DNA.



Supplementary Figure 4. a. A sample 10 kbp YOYO-1 event has several parameters of data analysis illustrated. Peak intensity was computed as the difference between average baseline fluorescence and peak level. Dwell time was the time between initial spike and the drop below a chosen threshold. Each event was fit with an exponential. Photobleaching time was reported as the time constant of the fitted curve, 803 ± 6 ms in this example. b. Histogram of dwell times and an exponential fit for 1 kbp data.

5. YOYO-stained DNA fluorescence intensity model

We model the dependence of intensity vs. length by considering the anticipated signal from a fully packed DNA molecule inside the NZMW. First, the illumination intensity decays exponentially with height *z* according to $I(z) = I_0 exp(-z/\Lambda)$, where Λ is a constant determined by the optical properties of the system (metal type, hole size, coating, illumination wavelength)[2]. The total anticipated fluorescence intensity, *F*, can be expressed as:

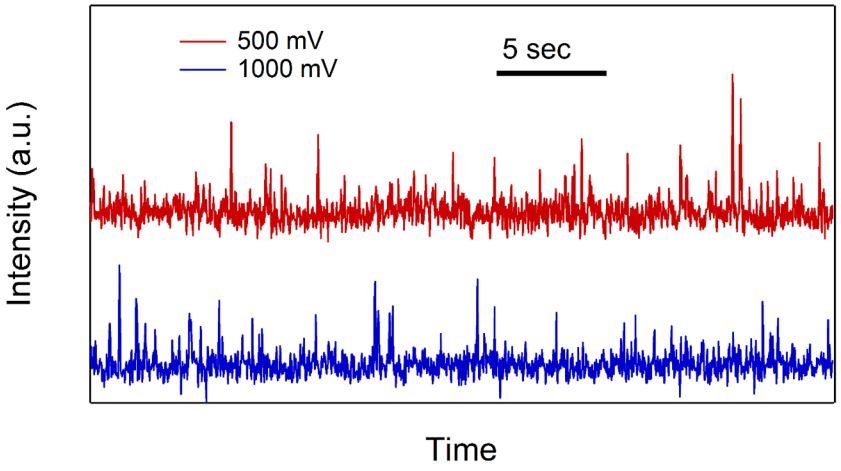
$$F(h) = F_0 \int_0^h e^{-z/\Lambda} dz \qquad (1)$$

where F_0 is the maximum fluorescence for a uniform-density, fully packed NZMW and *h* is the filled NZMW height. If the voltage-induced electric field gradient packs a DNA molecule inside the NZMW from floor to height *h*, then substituting *h* for "packed DNA height" *bN* (b = packed DNA z-height per kbp, N = DNA length in kbp), and solving the integral in **Eqn. 1**, yields:

$$F(N) = A\left(1 - e^{-bN/\Lambda}\right) \qquad (2)$$

6. Ribosomal RNA capture

As sample RNA, we used RNA from the 70S ribosome of *E. coli* (Sigma). To fluorescently label the molecule for optical NZMW detection, we incubated RNA with SYBR Green II (ThermoFisher) using the manufacturer's standard protocol. In Figure S5, we plot extracted traces for rRNA capture into an NZMW at 500 and 1000 mV bias.



Supplementary Figure 5. Ribosomal RNA molecules labeled with SYBR Green II produce fluorescent spikes when they enter a biased NZMW. The figure plots extracted fluorescence traces from an NZMW biased at 500 mV and 1000 mV.

7. Estimate of biotin-streptavidin binding time

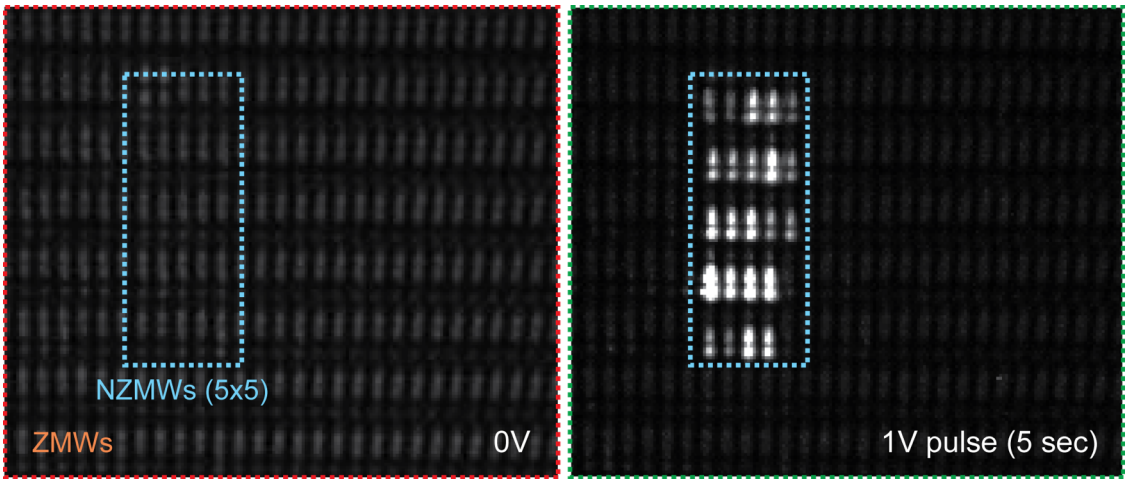
To estimate the necessary binding time for a streptavidin-conjugated sample to a biotin at the ZMW base, we modeled the reaction as a pseudo-first order process, with the concentrations of biotin ([B]), streptavidin ([S]), and biotin-streptavidin conjugate ([BS]) obeying the following equation.

$$\frac{d[BS]}{dt} = k[B][S]$$

The volume of a 100 nm diameter ZMW is 7.9 x 10^{-19} liters. Having one molecule (1.7 x 10^{-24} moles) in this volume corresponds to a ~μM concentration of ~μM. Using measured biotin-streptavidin association rate of 4.5 x 10^{7} M^{-1} s^{-1} for the first binding site of a streptavidin[3], the above equation suggests that the amount of time necessary to produce a 2 μM concentration of BS is roughly:

$$(1 \times 10^{-6}\ M) / (4.5 \times 10^{7}\ M^{-1}\ s^{-1} \times 1 \times 10^{-6}\ M \times 1 \times 10^{-6}\ M) \approx 0.022\ s.$$

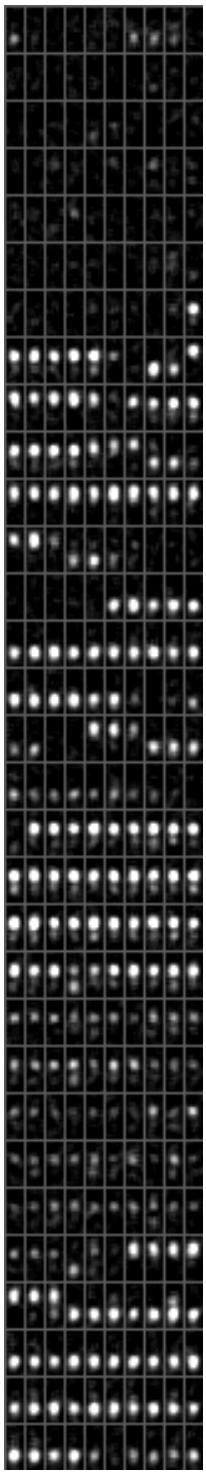
8. NZMW array loading at 1V



Supplementary Figure 6. Voltage loading YOYO-1 stained DNA into a 5x5 NZMW array using two 5 s, 1V pulses, results in 92% efficiency ([DNA] = 96 pg/μl).

After two subsequent 1 V pulses of 5 seconds, we loaded YOYO-1 stained, biotinylated, 48.5 kbp DNA into 23 out of 25 NZMWs in an array. In Figure S6, we present an image of the array before voltage pulse and an integrated stack of images during and after the pulses. Presence of immobilized DNA in a ZMW is verified by the fluorescence from the waveguide, which persisted after voltage was changed to 0 V.

9. Raw data and extracted sequence



Range 1: 7327 to 7353 Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
21.1 bits(13)	0.59	20/27(74%)	0/27(0%)	Plus/Plus

```
Query  35    ACAGTGTGCTCAGCCCTGACAACACCA  61
             ||| |||  ||||  |  |  ||||||||
Sbjct  7327  ACAATGTCTTCAGGCATAGCAACACCA  7353
```

Supplementary Figure 7. Sequence of frames (top to bottom, left to right) acquired for the 20 kbp SMRTbell template, as well as a read that corresponds to the sequence. Movie that shows this sequence of frames is also in the Supplementary Information.

10. Sequencing Details and base-calling algorithm

1) Analysis details

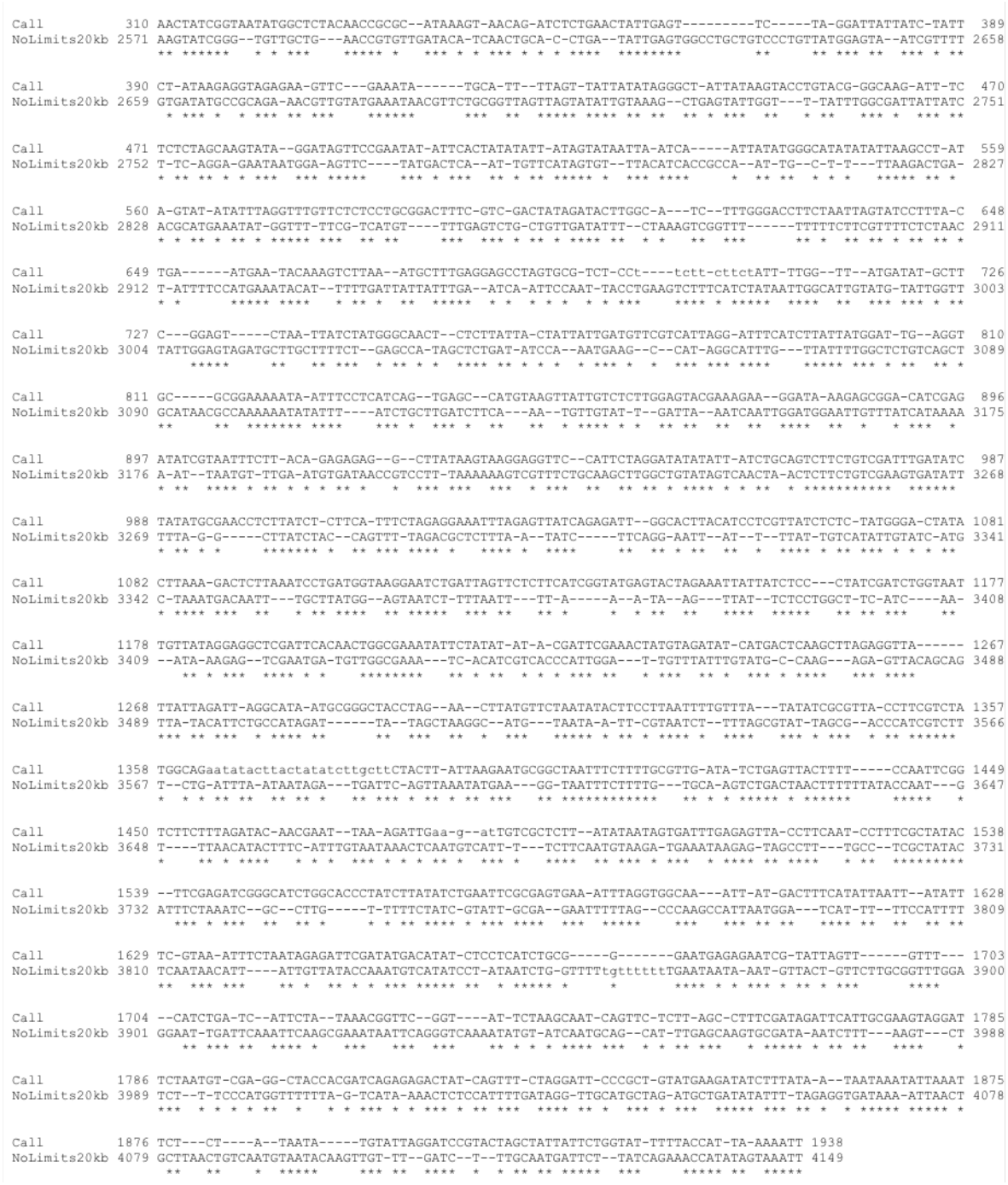
a. **Pulse sensing:** Our base-calling algorithm was carried out in the following 3-step process: The active ZMW is first located and two regions of interest are defined for the top and bottom portions of the spectra, which correspond to the G, T and C, A bases respectively. A threshold based algorithm adapted from the Pyth-Ion software written in the Wanunu lab *(https://github.com/rhenley/Pyth-Ion/)* was then run on each of the signals produced by the two regions of interest to identify start and end times for fluorescence pulses produced in each region. The two signals are then compared against one another to decouple their signals and ensure only one region is deemed as incorporating at any given time.

b. **Inter-pulse sensing:** In the event where multiple pulses appear as a single pulse due to the framerate being too slow to detect the inter-pulse time, a peak detection algorithm is then run to look for sudden changes in fluorescence intensity within pulses. If multiple peaks are detected then the pulse is separated into multiple pulses.

c. **Base Calling:** Finally, the frames from each of the detected pulses are combined by taking the median value of each pixel during the pulse. This composite image is then categorized by taking the dot product of the image with the dNTP standard images. Depending upon if the pulse comes from the top half or bottom half of the image, it will be assessed against either the dGTP and dTTP controls or the dCTP and dATP images respectively. The identity of the base is then determined based upon which control image gives the highest value.

2) Sequence alignment

a. **LAST Alignment:** Sequences were aligned using the LAST alignment software.[4] Sequence alignments of up to 1.6 kbp were achieved, producing sequence identity of 67% from a single read. Obtaining sequence identity on par with Pac Bio accuracy levels (85-90%) will require further mitigation of intrinsic SiN photoluminescence and further software refinement, but is beyond the scope of this work. An example of one read is shown below (obtained from a 10-minute read):



References

1. Assad, O. N.; Di Fiori, N.; Squires, A. H.; Meller, A., Two Color DNA Barcode Detection in Photoluminescence Suppressed Silicon Nitride Nanopores. *Nano Lett.* **2015,** *15* (1), 745-752.
2. Levene, M. J.; Korlach, J.; Turner, S. W.; Foquet, M.; Craighead, H. G.; Webb, W. W., Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* **2003,** *299* (5607), 682-686.
3. Srisa-Art, M.; Dyson, E. C.; deMello, A. J.; Edel, J. B., Monitoring of Real-Time Streptavidin−Biotin Binding Kinetics Using Droplet Microfluidics. *Anal. Chem.* **2008,** *80* (18), 7063-7067.

4. Kiełbasa, Szymon M., et al. Adaptive seeds tame genomic sequence comparison. *Genome research* **2011**, 21, 487-493.