iScience, Volume *23*

Supplemental Information

Direct microRNA Sequencing Using Nanopore-Induced Phase-Shift Sequencing

Jinyue Zhang, Shuanghong Yan, Le Chang, Weiming Guo, Yuqin Wang, Yu Wang, Panke Zhang, Hong-Yuan Chen, and Shuo Huang

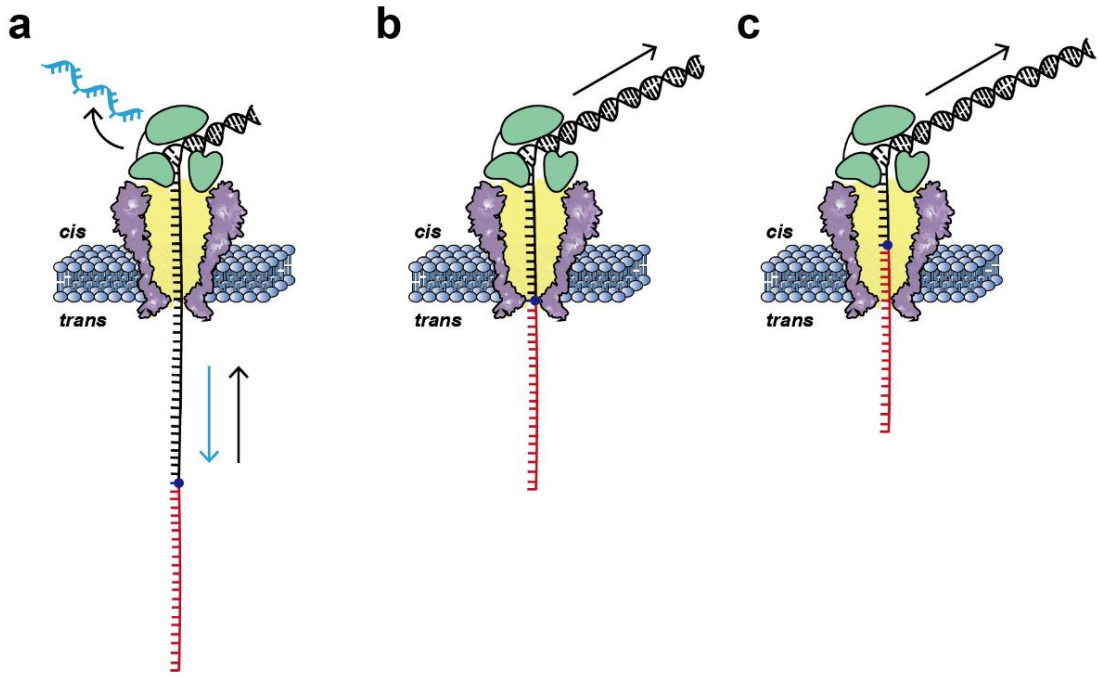

Figure S1: Detailed schematic diagrams of NIPSS, Related to Figure 1. The configuration of direct miRNA sequencing by NIPSS is as reported previously (Yan et al., 2019). The chimeric template is composed of a miRNA segment (red), an abasic residue (blue dot) and a DNA segment (black) **(Methods 1)**. With a +180 mV applied potential, the sequencing library, which was bound with a phi29 DNAP (green), was driven electrophoretically into the MspA nanopore (purple). **a)** The initiation of NIPSS. The cyan DNA blocker strand, which was thermally annealed with the sequencing library, was first voltage-driven unzipped from the chimeric template. This unzipping consequently triggered the replication-driven ratcheting by the phi29 DNA polymerase on top of the nanopore. The transition from voltage-driven unzipping to replication-driven ratcheting represents the initiation of nanopore sequencing. The motion directions of the chimeric template during unzipping and ratcheting are marked with cyan and black arrows, respectively. **b)** The initiation of miRNA sequencing. Passage of the abasic spacer through the pore constriction marks the initiation of subsequent miRNA sequencing. **c)** MiRNA sequencing by NIPSS. By utilizing the phase-shift, the miRNA segment passes through the pore constriction with single nucleotide steps, when the DNA drive-strand is being replicated by the phi29 DNA polymerase. All NIPSS experiments in this paper follow this configuration. All sequencing experiments were performed at 23 ^{o}C in the sequencing buffer of 0.3 M KCl, 10 mM HEPES/KOH, 10 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$ and 4 mM DTT at pH 7.5.

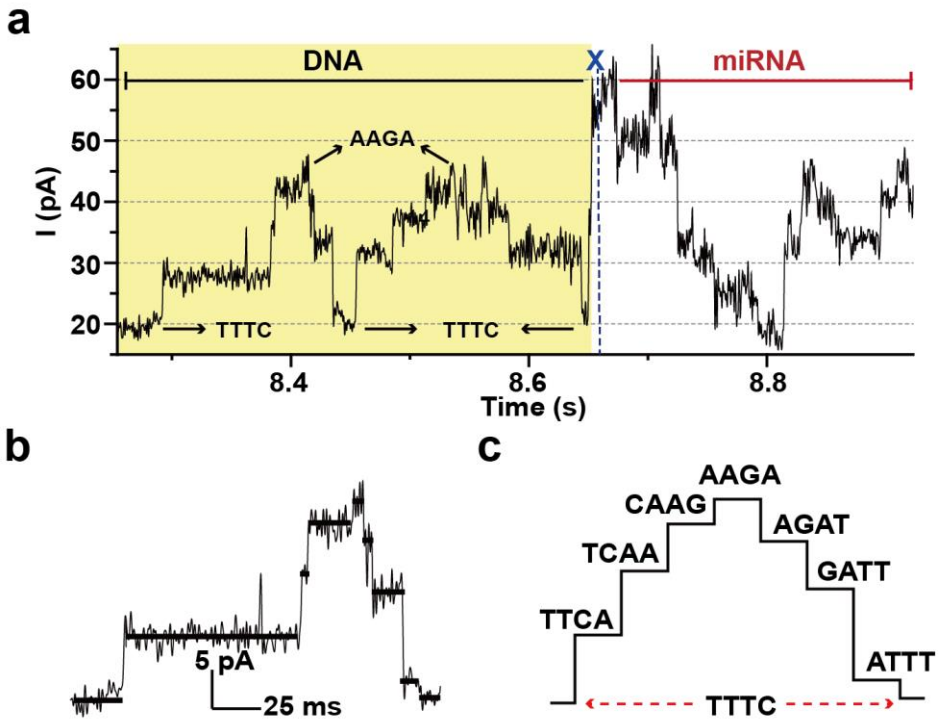

Figure S2: Design of the DNA segment within the chimeric template, Related to Figure 1. **a)** A representative current trace acquired by sequencing DNA-miR-21 using NIPSS. The trace segments that correspond to reading DNA, the abasic spacer and miRNA are marked appropriately. The purpose of including the DNA segment, which is designed on the 5'-end of the chimeric template strand **(Table S1)**, is twofold. First, it acts as a drive strand, which can be enzymatically ratcheted by the phi29 DNAP against the electrophoretic force, guarantees that the miRNA segment can be sequenced by nanopores. Second, by including two sequence repeats of "AGAACTTT" (5'-3') in the DNA segment **(Table S1)**, a unique trace pattern of two triangular wave (region marked with yellow shade) will appear ahead of the miRNA reading, which helps to recognize the initiation of a NIPSS event. Nanopore reading of AAGA and TTTC (3'-5') marked by arrows generates the highest and the lowest residual current among that from all other canonical combinations of DNA quadromers respectively. **b)** The characteristic current signature generated by the DNA segment. The demonstrated trace was extracted from the first repetitive cycle of DNA segment in **a**. Scale bar: 5 pA/25 ms. **c)** The extracted current steps from **b** with aligned quadromer sequences. The sequencing steps correspond to different quadromer readings during NIPSS, in which AAGA and TTTC (3'-5') show the highest and lowest current step, respectively.

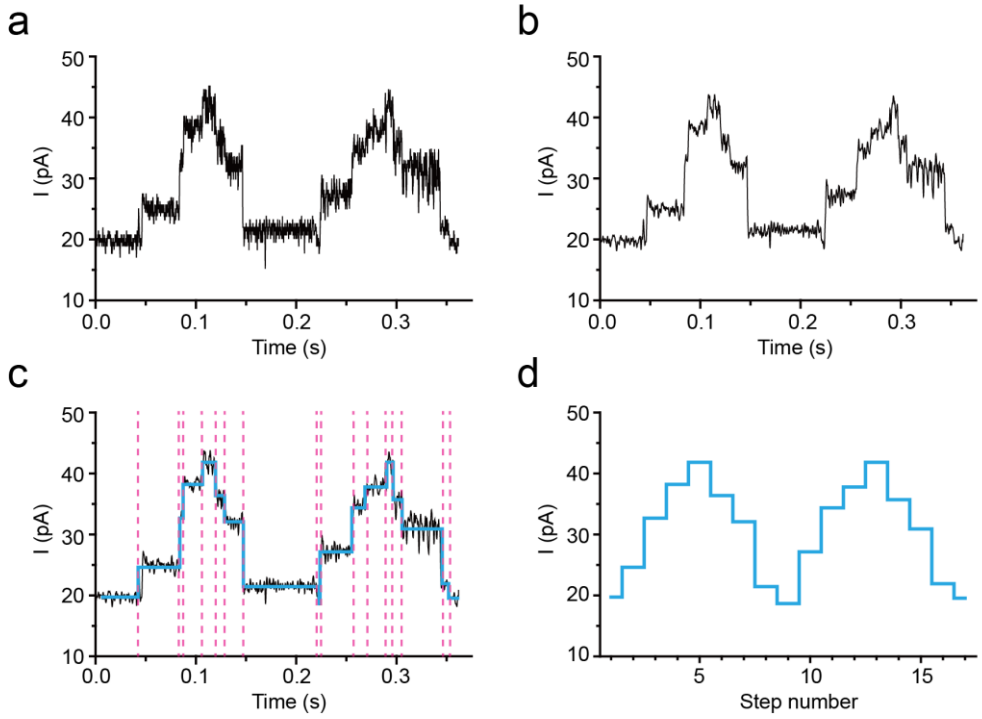

Figure S3: Squiggle plot generation, Related to Figure 1. a) A demonstrative raw nanopore sequencing data. The trace was acquired by a Digidata 1550B digitizer (Molecular Devices) with a 25 kHz sampling rate and low-pass filtered at 5 kHz. **b)** Post-acquisition signal processing. The raw sequencing trace in **a** was further digitally filtered by a low pass Gaussian filtered with a 500 Hz cutoff frequency. Noise at high frequency is significantly reduced after this step. Step transitions resulted from nanopore sequencing is clearly visible in the trace. **c)** Trace segmentation. Step transitions within the trace were automatically detected, as marked with pink dashed lines. Mean values of event steps were separately derived, as marked with blue lines. **d)** Squiggle plot. The squiggle plot was generated by plotting mean values of each event step according to their order of appearance.

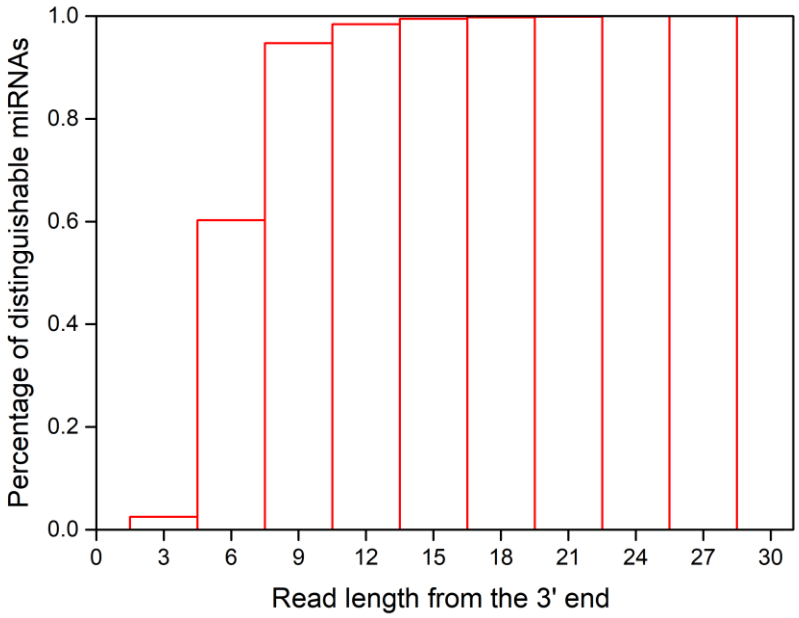

Figure S4: The distinguishable human miRNAs by NIPSS, Related to Figure 1. The sequences of all identified mature human miRNAs were downloaded from miRbase (http://www.mirbase.org/) to form an archive for follow-up analysis. By the time this article is written, this archive contains 2656 human mature miRNAs. It should be noted that some miRNAs share completely identical bases in this archive. For example, hsa-miR-199a-3p and hsa-miR-199b-3p has an identical sequence. These miRNAs originate from either distinct pre-miRNAs or identical pre-miRNAs found in different genome loci (Zhong et al., 2019). By excluding these types of miRNAs with identical sequences, 23 miRNAs have been deleted from the archive, forming an archive of 2633 miRNAs, each with a unique sequence. By reading a different length of nucleotides from their 3' ends, the percentage of distinguishable miRNAs were evaluated. Here, a distinguishable miRNA is counted when this miRNA sequence with a particular read length from their 3' end is different from any other miRNAs in the archive. According to this test, it has been evidenced that by reading the first 15 nucleotides from their 3' end, 99.5% of human miRNA become distinguishable **(Table S2)**.

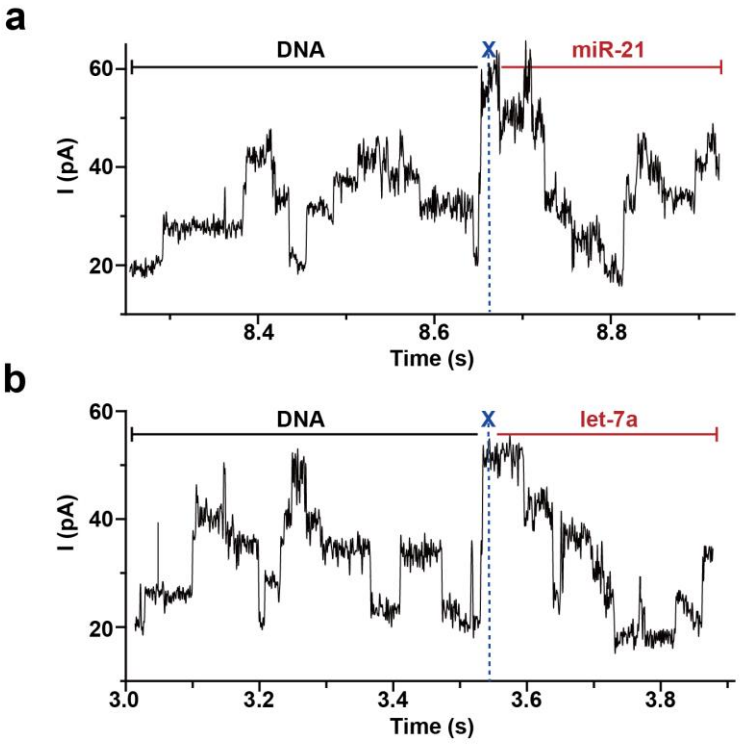

Figure S5: Identification of miR-21 and let-7a by NIPSS, Related to Figure 2. MiRNA identities can be directly recognized by analyzing the miRNA part of the NIPSS signal. **a)** A representative current trace from DNA-miR-21 sequencing. **b)** A representative current trace from DNA-let-7a sequencing. The initiation of miRNA sequencing is indicated by blue dashed lines, where an abasic spacer (X) is read by the nanopore. The signal patterns of the DNA segment show high similarities since the sequence of the DNA drive-strand is identical. Whereas, the miRNA segment of the signal shows remarkable differences between the demonstrated NIPSS events.

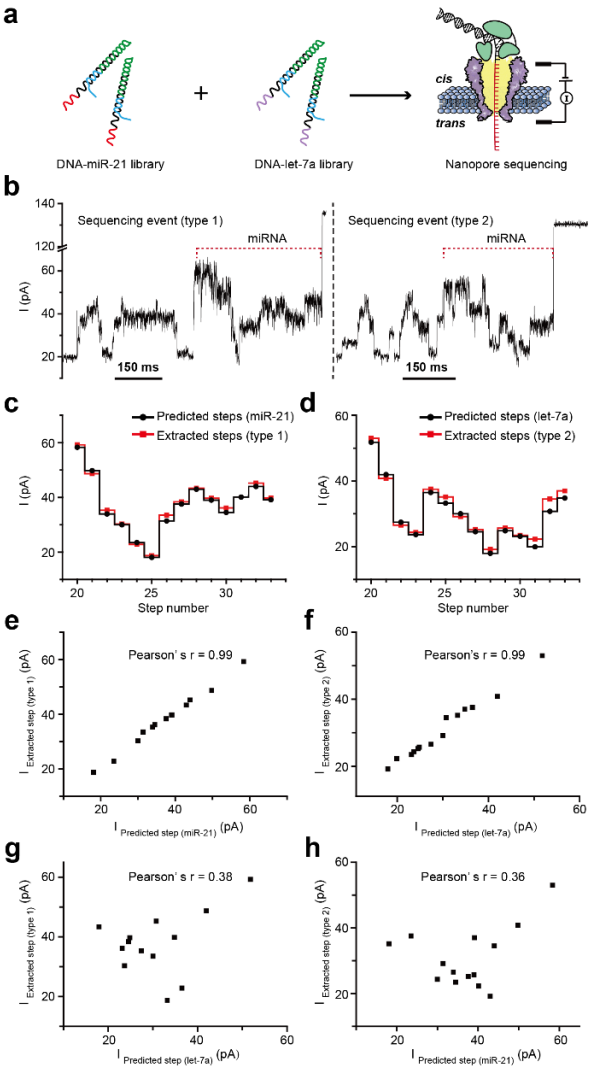

Figure S6: Discrimination of DNA-miR-21 and DNA-let-7a from a mixture, Related to Figure 2. a) The schematic diagram of NIPSS miRNA sequencing from a mixture of analyte. The sequencing library complex was respectively prepared from DNA-miR-21 and DNA-let-7a with a 5 nM final concentration for each composite. **b)** Representative raw sequencing traces acquired from a mixture of DNA-miR-21 and DNA-let-7a measured by the same nanopore**.** Two discriminable types of sequencing events (type 1 and 2), which differs in their miRNA segments (red dashed lines), were clearly recognized. Steps of transition in miRNA segments of each event were extracted as described in **Fig. S3** to form squiggle plots marked as extracted steps in **c** and **d**. The predicted steps of miR-21 and let-7a were respectively created from mean values of sequencing steps acquired from multiple independent miR-21 and let-7a events as demonstrated in **Fig. 2c**. The extracted steps from type 1 or type 2 respectively show strong resembalance with predicted steps from miR-21 or let-7a. To quantitatively evaluate their correlation, results in **c** and **d** were respectively plotted in **e** and **f**, from which the extracted value of one type of event and their predicted values were taken as two sets of variables (x_i, y_i) indexed by i. Pearson correlation coefficient (Pearson's r) was calculated according to the equation $r = \sum_{i=1}^{n}(x_i - \bar{x})\,(y_i - \bar{y}) / \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}$, from which $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$ and of $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$. The Pearson's r between results of type 1 event and

miR-21 or between results of type 2 event and let-7a are both 0.99, indicating a strong linear correlation between. On the other side, as demonstrated in **g** and **h**, extremely weak correlations were observed between type 1 event and let-7a or between type 2 event and miR-21. The reported Pearson's r are for are 0.38 **(g)** and 0.36 **(h)**.

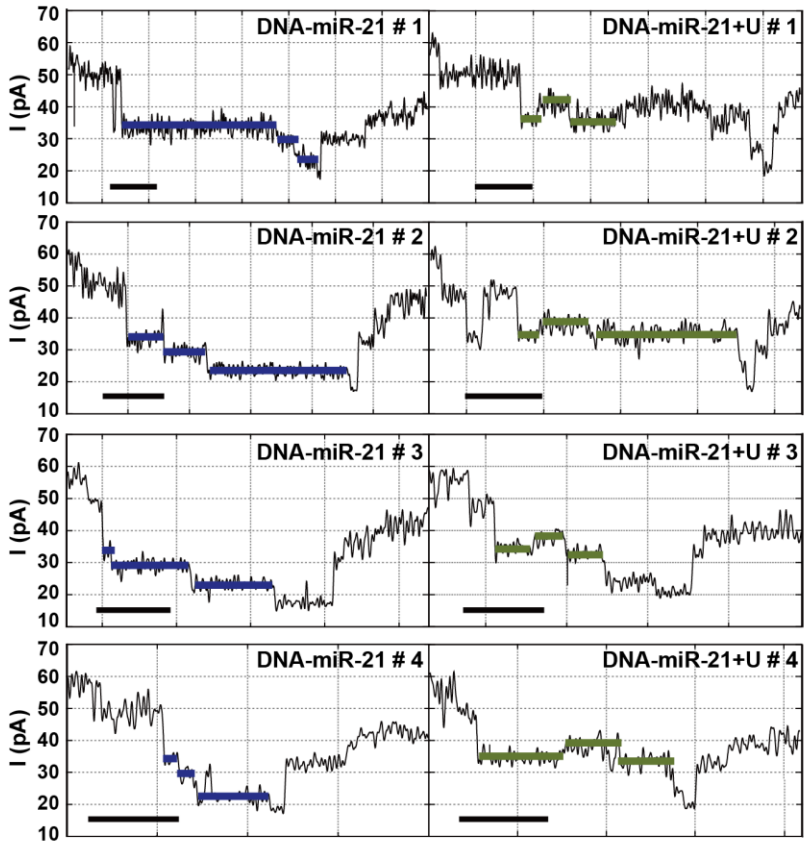

Figure S7: Representative current traces for isomiRs discrimination, Related to Figure 3. Each panel shows individual trace segments, which correspond to the miRNA part of the nanopore sequencing signal of DNA-miR-21 (left) and DNA-miR-21+U (right). Current steps with significant deviations between analytes are indicated by solid lines (blue for miR-21 and yellow for miR-21+U). The duration time for each nanopore sequencing step is stochastic. Scale bar: 50 ms.

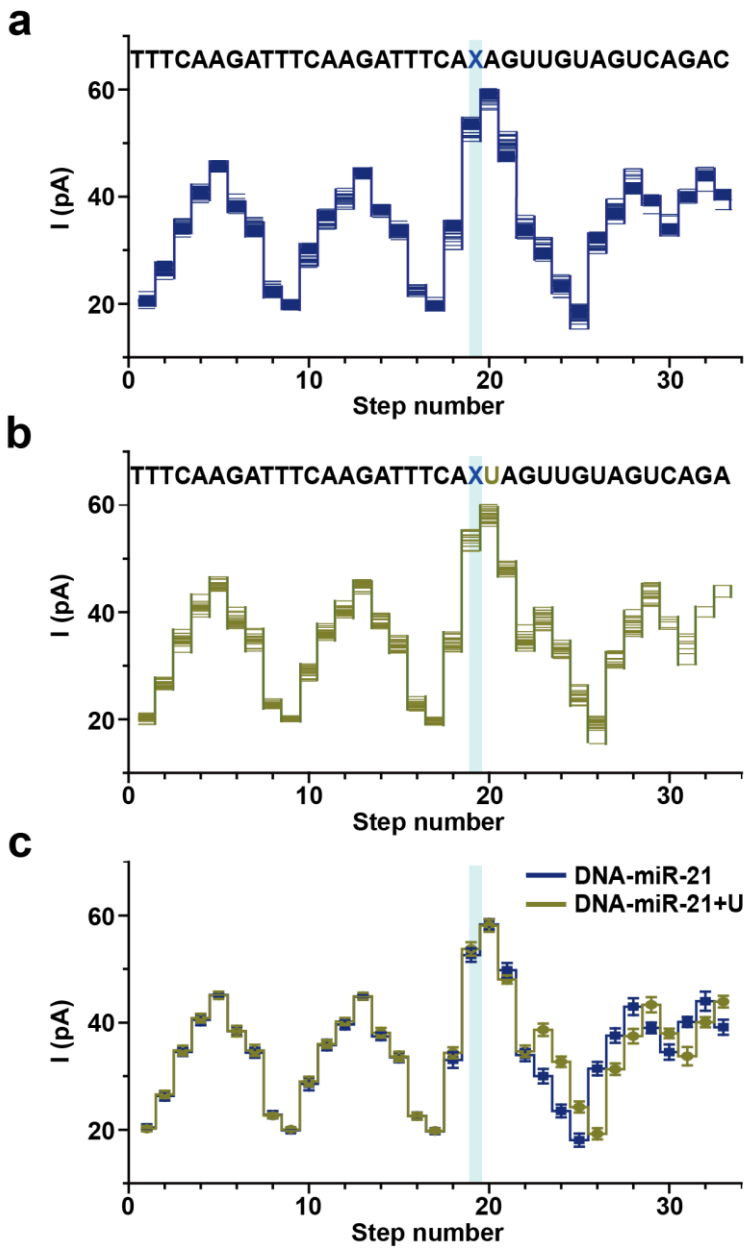

Figure S8: Statistical comparison between DNA-miR-21 and DNA-miR-21+U, Related to Figure 3. a) Overlay of extracted current steps from multiple DNA-miR-21 events (N=24). **b)** Overlay of extracted current steps from multiple DNA-miR-21+U events (N=24). The correlated sequences for each sequence are demonstrated on top of the plots in **(a-b)** where an abasic site (X) is located between DNA and miRNA. **c)** Consensus comparison of sequencing signals between DNA-miR-21 and DNA-miR-21+U. Mean and standard deviations extracted from **a** and **b**. Here, a uridine monophosphate insertion has generated an additional current step when reading DNA-miR-21+U in reference to that of DNA-miR-21. Consequently, the signal pattern from DNA-miR-21+U is shifted by 1 nucleotide. The blue strip in **(a-c)** represents the quadromer reading of TCAX, which is the first quadromer containing the abasic site read by NIPSS.

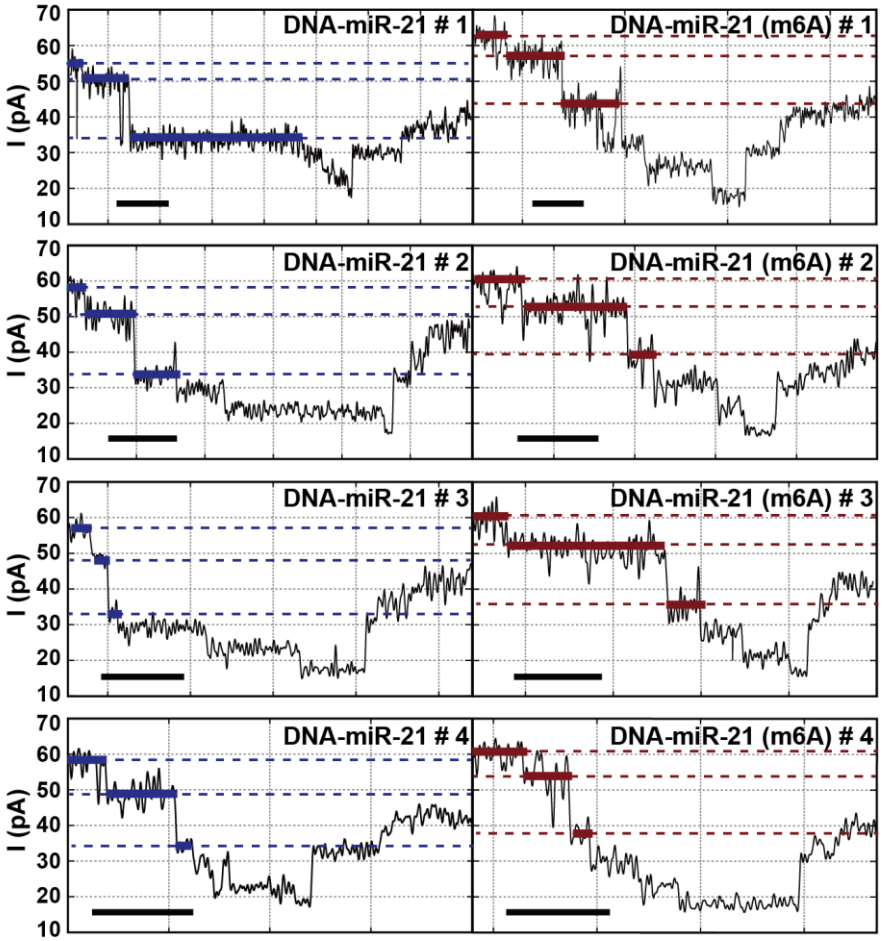

Figure S9: Representative current traces for m6A modification detection, Related to Figure 4. The sequence contexts between DNA-miR-21 and DNA-miR-21(m6A) differ with only one m6A modification **(Table S1)**. Each panel shows individual trace segments, which correspond to the miRNA part of the sequencing signal of DNA-miR-21 (left) and DNA-miR-21(m6A) (right). Current steps with significant deviations between analytes are indicated by solid lines. These level differences could also be recognized from the dashed reference lines. The duration time for each nanopore sequencing step is stochastic. Scale bar: 50 ms.

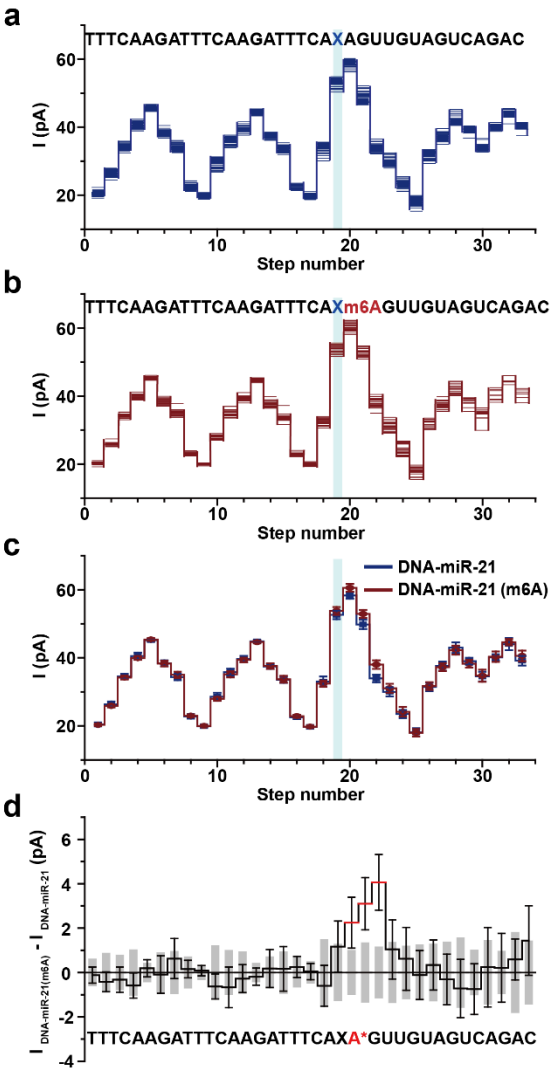

Figure S10: Statistical signal variation between DNA-miR-21 and DNA-miR-21(m6A), Related to Figure 4. a) Overlay of current steps extracted from multiple DNA-miR-21 events (N=24). **b)** Overlay of current steps extracted from multiple DNA-miR-21(m6A) events (N=24). Each current step in **(a-b)** represents the mean current values of each quadromer reading. The correlated sequences are shown above the plots. The "X" within the sequence stands for the abasic site located between the DNA and the miRNA segment of the chimeric template. **c)** Demonstration of different current patterns generated from two DNA-miRNA strands. The average current values and error bars are created from current level traces exhibited in **a** and **b**. The blue strips in **(a-c)** represent the quadromer reading of TCAX, which is the first quadromer containing the abasic site when read by NIPSS. **d)** Statistical current differences between DNA-miR-21 and DNA-miR-21(m6A). Mean current values in **c** were used to construct the current difference map, where the value of $I_{DNA-miR-21(m6A)} - I_{DNA-miR-21}$ is displayed by a black solid line. The standard deviation of signals from DNA-miR-21 is shown with gray columns. The standard deviation of signals from DNA-miR-21(m6A) is demonstrated with black bars. Signal variations caused by m6A modification are shown as three red lines. The sequence of the strand is aligned below the figure, where A* (red) represents either A or m6A in the sequence.

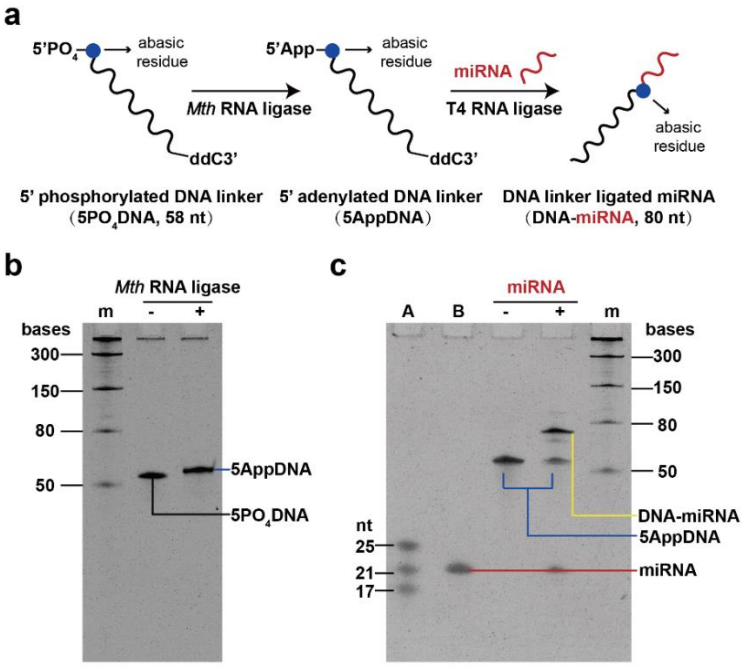

Figure S11: Enzymatic ligation between DNA and RNA, Related to Figure 1. a) A schematic diagram of enzymatic ligation. The reaction starts with a 58 nt DNA linker ($5PO_4$DNA), which has a dideoxycytosine (ddC) on its 3' end and a phosphorylated abasic residue on its 5' end. The abasic site is indicated by blue dot. After treatment with the *M*th RNA ligase (New England Biolabs), the 5' end of the DNA linker is adenylated (5AppDNA). To minimize non-specific ligation, T4 RNA ligase 2 truncated K227Q mutant (T4 Rnl2tr, New England Biolabs) was chosen, which specifically ligates the 5' end of the pre-adenylated DNA linker (5AppDNA) with the 3' end of target miRNA. **b)** Characterization of DNA adenylation by 15% polyacrylamide (PAGE)-urea gel. The adenylation reaction was performed by mixing the following components: 200 pmol $5PO_4$DNA, 4 µL 10x 5' DNA adenylation buffer, 4 µL 1 mM ATP, 4 µL *M*th RNA liagse and nuclease-free H_2O to a final volume of 40 µL. The reaction was incubated at 65 ^oC for 1 h and heat inactivated by incubation at 85 ^oC for 5 min. The reaction product 5AppDNA was purified by ethanol precipitation. Lane + and lane – stands for samples that were incubated with *M*th RNA ligase or not, respectively. Lane m stands for the Low Range ssRNA Ladder (New England Biolabs). 5' adenylation results in slight upshift of the band during gel electrophoresis. **c)** Characterization of DNA-miRNA ligation. Enzymatic ligation results were characterized by 15% PAGE-Urea gel electrophoresis. The ligation reaction was performed by mixing the following components: 5.0 µL 50% (w/v) PEG 8000, 2 µL 10x RNA ligase buffer, 20 pmol 5AppDNA, 10 pmol miRNA, 1 µL T4 Rnl2tr, nuclease-free water to a final volume of 20 µL. The reaction was incubated at 4 °C for 24 h and heat inactivated by incubation at 65 °C for 20 min. Lane A indicates the microRNA Marker (New England Biolabs). Lane B stands for the miR-21 strand. Lane - stands for the 5AppDNA. Lane + stands for the ligation product. Lane m stands for the Low Range ssRNA Ladder (New England Biolabs). An extra, high molecular weight band was detected in lane m, which is the ligated DNA-miRNA chimeric strand.

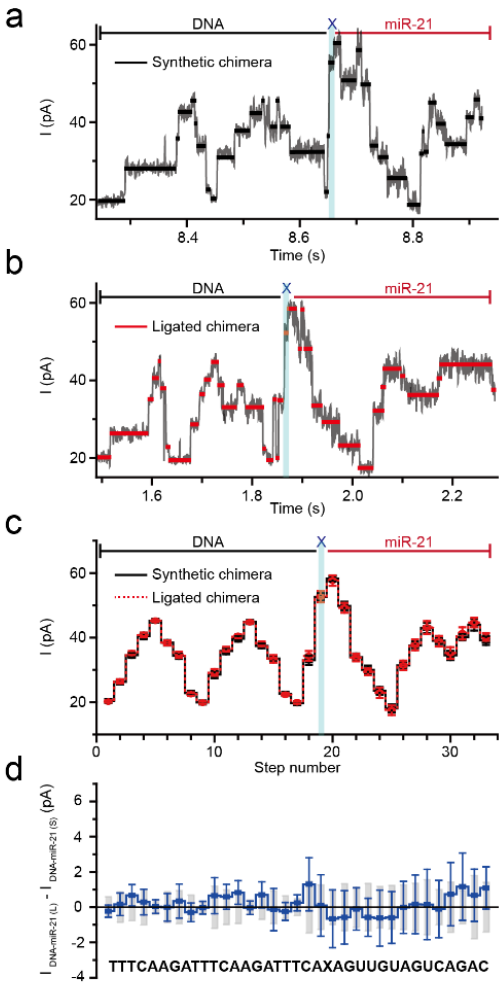

Figure S12: NIPSS sequencing of DNA-miR-21 from the ligated chimera, Related to Figure 1. a) A representative raw trace from NIPSS sequencing of the synthetic DNA-miR-21 chimera. **b)** A representative raw trace from NIPSS sequencing of the ligated DNA-miR-21 chimera. The DNA and miRNA segments of the trace were respectively marked with black and red lines. The trace segment of reading the abasic residue (X) was marked with a blue stripe. Solid lines over the traces in **(a-b)** represent the extracted mean current values of the signal plateau. The demonstrated results were acquired with an aqueous buffer of 0.3 M KCl, 10 mM HEPES/KOH, 10 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$ and 4 mM DTT at pH 7.5. **c)** An overlay of current steps between two DNA-miR-21 chimeras. Results in the squiggle plots were respectively acquired from either the synthetic or the ligated DNA-miR-21 chimera. The mean and standard deviations values were respectively created from 20 independent sequencing events similar to those shown in **a** and **b**. All steps show a strong resemblance, indicating that the enzymatically ligated chimera was successfully sequenced by the NIPSS configuration. **d)** Statistical current differences between the synthetic DNA-miR-21 chimera and the ligated DNA-miR-21 chimera. Mean current values in **c** were used to construct the current difference map, where the value of $I_{DNA-miR-21(L)} - I_{DNA-miR-21(S)}$ is displayed by a blue solid line. Here, the "L" or "S" label respectively represents the ligated chimera or the synthetic chimera. The standard deviation values from the synthetic DNA-miR-21 chimera were shown with gray columns. The standard deviation values from the ligated DNA-miR-21 chimera were

demonstrated with blue bars. The mean current differences appear negligible between the two sources of DNA-miR-21 chimeras, indicating that the enzymatic ligation described in **Fig. S11** could be used to construct NIPSS sequencing libraries from natural miRNAs.

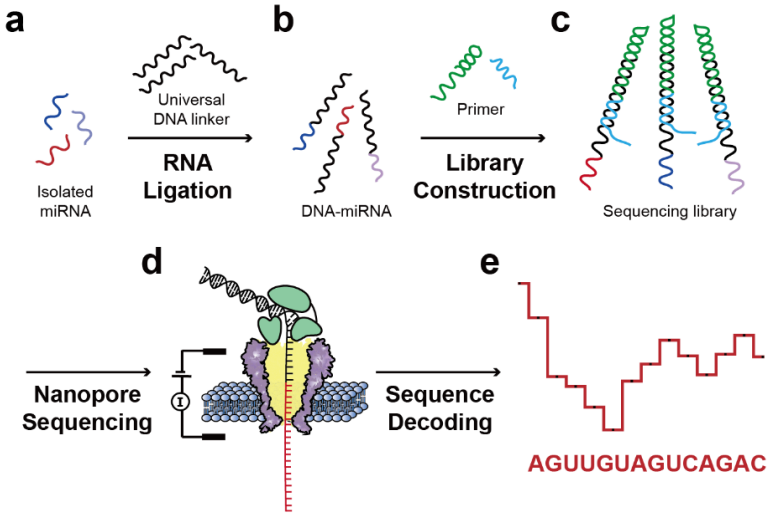

Figure S13: A proposed strategy to directly sequence miRNA from natural resources, Related to Figure 1. a-c) Schematic diagram of NIPSS sequencing library preparation with isolated miRNA from natural resources. Isolated miRNAs **a)** could be ligated to form DNA-miRNA chimeric strands **b)** and subsequently form sequencing libraries **c)** by thermal annealing. **d)** Direct miRNA sequencing is carried out as described in this paper. **e)** Current step transitions during NIPSS reading of miRNA could be decoded into RNA sequences for downstream clinical diagnosis or bioinformatics investigations.

Table S1: Nucleic acid sequences used in this study, Related to Figure 1, 2, 3 and 4, and to Figure S11 and S12.

Construct	Sequence (5'-3')
DNA-miR-21	UAGCUUAUCAGACUGAUGUUGAXACTTTAGAACTTTAGAACTTTTCAGATCTC ACTATCGCATTCTCATGCAGGTCGTAGC
DNA-let-7a	UGAGGUAGUAGGUUGUAUAGUUXACTTTAGAACTTTAGAACTTTTCAGATCTC ACTATCGCATTCTCATGCAGGTCGTAGC
DNA-miR-21+U	UAGCUUAUCAGACUGAUGUUGAUXACTTTAGAACTTTAGAACTTTTCAGATCT CACTATCGCATTCTCATGCAGGTCGTAGC
DNA-miR-21 (m6A)	UAGCUUAUCAGACUGAUGUUGm6AXACTTTAGAACTTTAGAACTTTTCAGATC TCACTATCGCATTCTCATGCAGGTCGTAGC
Primer	GCGTACGCCTACGGTTTTCCGTAGGCGTACGCGCTACGACCTGCATGAGAATG C
Blocker	GATAGTGAGATCTGATTTCCCAAATTTAAA-cholesterol
DNA linker	PO_4- XACTTTAGAACTTTAGAACTTTTCAGATCTCACTATCGCATTCTCATGCAGGT CGTAG-dideoxyC
miR-21	UAGCUUAUCAGACUGAUGUUGA

Notes:

1. All miRNA segments of the sequence contexts are underlined.
2. Primer/Template duplex regions in the chimeric template strand are indicated by blue letters.
3. Blocker/Template duplex regions in the chimeric template strand are indicated by red letters.
4. The "X" letter represents the abasic site.
5. PO_4 stands for the 5' phosphate group.
6. DideoxyC stands for a 3' dideoxycytosine.

Table S2: The values of distinguishable ratio for various read length from 3' end of miRNAs, Related to section "Direct miRNA sequencing using NIPSS" in main-text, and to Figure 1.

RNA read length from the 3' end	Counts of undistinguishable miRNAs	Counts of distinguishable miRNAs	Distinguishable ratio (%)
3	2567	66	2.51
6	1045	1588	60.31
9	138	2495	94.76
12	41	2592	98.44
15	14	2619	99.47
18	6	2627	99.77
21	2	2631	99.92
24	0	2633	100
27	0	2633	100

Table S3: Current values of all 4-nucleotides sequence contexts for isomiRs, Related to Figure 3.

Level number	miR-21			miR-21+U		
	Sequence quadromers (3'-5')	Mean I value (pA)	Standard Deviation (pA)	Sequence quadromers (3'-5')	Mean I value (pA)	Standard Deviation (pA)
1	TTTC	20.4	0.6	TTTC	20.2	0.5
2	TTCA	26.3	0.9	TTCA	26.6	0.7
3	TCAA	34.5	0.8	TCAA	34.8	0.9
4	CAAG	40.5	1.0	CAAG	40.9	0.8
5	AAGA	45.2	0.4	AAGA	45.1	0.7
6	AGAT	38.4	0.9	AGAT	38.4	1.0
7	GATT	34.4	0.9	GATT	34.8	1.0
8	ATTT	22.8	0.7	ATTT	22.6	0.5
9	TTTC	19.9	0.3	TTTC	20.0	0.3
10	TTCA	28.6	1.2	TTCA	29.0	0.9
11	TCAA	35.8	1.0	TCAA	36.0	0.7
12	CAAG	39.7	0.9	CAAG	40.2	0.7
13	AAGA	44.8	0.4	AAGA	44.9	0.7
14	AGAT	37.4	0.7	AGAT	38.1	0.9
15	GATT	33.5	0.9	GATT	33.7	0.9
16	ATTT	22.6	0.6	ATTT	22.5	0.7
17	TTTC	19.7	0.5	TTTC	19.8	0.4
18	TTCA	33.0	1.5	TTCA	34.4	1.0
19	TCAX	52.6	1.3	TCAX	53.7	1.3
20	CAXA	58.3	1.0	CAXU	58.1	1.1
21	AXAG	49.8	1.3	AXUA	48.0	0.8
22	XAGU	33.9	1.1	XUAG	34.7	1.1
23	AGUU	30.0	1.3	UAGU	38.7	1.2
24	GUUG	23.5	1.2	AGUU	32.7	1.0
25	UUGU	18.1	1.2	GUUG	24.2	1.0
26	UGUA	31.4	1.3	UUGU	19.2	1.0
27	GUAG	37.6	1.4	UGUA	31.3	1.1
28	UAGU	43.0	1.6	GUAG	37.5	1.4
29	AGUC	39.0	1.0	UAGU	43.3	1.4
30	GUCA	34.5	1.4	AGUC	38.0	0.9
31	UCAG	40.1	0.9	GUCA	33.7	1.7
32	CAGA	44.0	1.8	UCAG	40.1	1.0
33	AGAC	39.1	1.4	CAGA	43.9	1.1

Transparent Methods

Materials

Potassium chloride (KCl), sodium chloride (NaCl), sodium hydrogen phosphate (Na_2HPO_4) and sodium dihydrogen phosphate (NaH_2PO_4) were obtained from Aladdin (China). Magnesium chloride ($MgCl_2$) was from Macklin. Ammonium sulfate $(NH_4)_2SO_4$ was from Xilong Scientific. 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES) was from Shanghai Yuanye Bio-Technology (China). Ammonium persulfate, kanamycin sulfate, dl-dithiothreitol (DTT), dioxane-free isopropyl-β-D-thiogalactopyranoside (IPTG), N,N,N',N'-Tetramethyl- ethylenediamine (TEMED) and imidazole were from Solarbio (China). Ethylene- diaminetetraacetic acid (EDTA), pentane, hexadecane and Genapol X-80 were from SIGMA-ALDRICH. 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC) was from Avanti Polar Lipids. Urea was from BIOSHARP. Acrylamide was from Sangon Biotech. The Low Range ssRNA Ladder (#N0364S), microRNA Marker (#N2102S), RNA loading dye (#B0363S), Nuclease-free Water (#B1500S), T4 RNA ligase 2 truncated K227Q mutant (T4 Rnl2tr; #M0242S), phi29 DNA Polymerase (#M0269S), deoxynucleotide (dNTP) solution mix (#N0447S) and the 5' DNA Adenylation kit (E2610S) were from New England Biolabs. *E. coli* strain BL21 (DE3) was from Biomed (China). Luria-Bertani (LB) agar and LB broth were from Hopebio (China).

All HPLC-purified DNA oligonucleotides, including DNA-miRNA template strands, primer, blocker, DNA linker and miR-21 **(Table S1)** were custom synthesized by Genscript (New Jersey, USA).

Methods

1. The construction of miRNA sequencing library.

The sequencing library for NIPSS is thermally annealed from three separate nucleic acid strands: the chimeric template, the primer and the blocker **(Fig. 1a, Table S1)**. These three strands were mixed with a 1:1:2 molar ratio in an aqueous buffer (0.3 M KCl, 10 mM HEPES/KOH, 10 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$). Thermal annealing was carried out by incubating the mixture at 95 °C for 2 min and program cooled down to 25 °C with a rate of -5 °C/min. For optimum sequencing data production, the thermal annealed sequencing library should be immediately used in subsequent NIPSS measurements.

2. The preparation of a biological nanopore.

The MspA mutant (D90N/D91N/D93N/D118R/D134R/E139K) nanopore (Butler et al., 2008) was expressed with *E. coli* BL21 (DE3) and purified with nickel affinity chromatography as described previously (Wang et al., 2018; Yan et al., 2019). This MspA mutant, which is the sole MspA nanopore discussed in this paper, is named MspA, if not otherwise stated. Briefly, the constructed plasmid gene for MspA was heat-shock transformed into *E. Coli* BL21 (DE3). Afterwards, the cells were grown in LB medium to an OD_{600}=0.7, induced with 1 mM IPTG and

shaken (180 rpm) overnight at 16 °C. The cells were harvested by centrifugation (4000 rpm, 20 min, 4 °C). The collected pellet was re-suspended in the lysis buffer (100 mM Na_2HPO_4/NaH_2PO_4, 0.1 mM EDTA, 150 mM NaCl, 0.5% (w/v) Genapol X-80, pH 6.5) and heated to 60 °C for 10 min. The suspension was cooled on ice for 10 min and centrifuged at 4 °C for 40 min at 13,000 rpm. After syringe filtration, the supernatant was loaded to a nickel affinity column (HisTrapTM HP, GE Healthcare). The column was first eluted with buffer A (0.5 M NaCl, 20 mM HEPES, 5 mM imidazole, 0.5% (w/v) Genapol X-80, pH=8.0) and further eluted with a linear gradient of imidazole (5 mM-500 mM) by mixing buffer A with buffer B (0.5 M NaCl, 20 mM HEPES, 500 mM imidazole, 0.5% (w/v) Genapol X-80, pH=8.0). The eluted fractions were further characterized by SDS-polyacrylamide gel electrophoresis (PAGE) and the desired protein was identified. The identified fraction was immediately used for NIPSS experiments or stored at -80 °C for long term storage.

3. NIPSS experiments.

NIPSS experiments were carried out as described previously (Yan et al., 2019). Briefly, the electrolyte buffer (0.3 M KCl,10 mM HEPES/KOH, 10 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$ and 4 mM DTT at pH 7.5) were separated by a 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC) lipid membrane (Avanti Polar Lipids) into *cis* and *trans* compartments. Both compartments were in contact with separate Ag/AgCl electrodes and connected to an Axopatch 200B patch clamp amplifier (Molecular Devices) to form a circuit, while the *cis* compartment is electrically grounded. Purified MspA nanopores were added in *cis* and spontaneously inserted into the membrane. With a single pore inserted **(Fig. 1b)**, the sequencing library, dNTPs and phi29 DNAP could be added into *cis* and stirred magnetically to reach final concentrations of 5 nM, 250 μM and 1 nM, respectively. Nanopore sequencing was initiated by holding an applied voltage at +180 mV. All electrophysiology recordings were acquired with a Digidata 1550B digitizer (Molecular Devices) with a 25 kHz sampling rate and low-pass filtered at 5 kHz. All NIPSS experiments were performed at room temperature (22 ±1 °C).

4. Data analysis.

Nanopore sequencing events were extracted from raw electrophysiology traces according to their unique pattern of signal transitions. The extracted events were low pass filtered at 500 Hz using Clampfit 10.7 (Molecular Devices) for further analysis **(Fig. S3)**. The sequencing steps, which represents signal plateau transitions within traces were fragmented. The average current for each step was calculated and finally extracted from sequencing events.

Supplemental References

Butler, T.Z., Pavlenok, M., Derrington, I.M., Niederweis, M., and Gundlach, J.H. (2008). Single-molecule DNA detection with an engineered MspA protein nanopore. Proc. Natl. Acad. Sci. USA *105*, 20647-20652.

Wang, Y., Yan, S., Zhang, P., Zeng, Z., Zhao, D., Wang, J., Chen, H., and Huang, S. (2018). Osmosis-Driven Motion-Type Modulation of Biological Nanopores for Parallel Optical Nucleic Acid Sensing. ACS Appl. Mater. Interfaces *10*, 7788-7797.

Zhong, X., Heinicke, F., and Rayner, S. (2019). miRBaseMiner, a tool for investigating miRBase content. RNA biology *16*, 1534-1546.