Pattern Recognition of microRNA Expression in Body Fluids using Nanopore Decoding at Sub-femtomolar Concentration

Nanami Takeuchi, Moe Hiratani, and Ryuji Kawano

Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Japan

Abstract

This paper describes nanopore decoding for microRNA (miRNA) expression patterns using DNA computing technology. miRNAs have shown promise as markers for cancer diagnosis due to their cancer type-specificity, and therefore simple strategies for miRNA-pattern recognition are required. We propose a system for pattern recognition of five types of miRNAs overexpressed in bile duct cancer (BDC). The information of miRNAs from BDC is encoded in diagnostic DNAs (dgDNAs) and decoded electrically by nanopore measurement. With this system, we succeeded in distinguishing miRNA expression patterns in the plasma of BDC patients using a label-free method and in real-time. Moreover, our dgDNA-miRNAs complexes can be captured by the nanopore at ultralow concentration, such as 0.1 fM. Such nanopore decoding with dgDNAs could be applied as a simple and early diagnostic tool for cancer in the future.

Introduction

DNA computing uses the biochemical reactions of information-encoding DNA molecules to solve problems. The autonomous calculations are implemented in-parallel in a wet

environment. This was proposed in the early pioneering work of Adleman in 1994, in which he presented a method to solve the Hamiltonian path problem - a mathematical problem also known as the traveling salesman problem – by using the DNA molecule itself to perform DNA computing.[1] After proposing this groundbreaking idea, extensive methods of DNA computation have been studied including logic gates.[2] The logic operation has been one of the most attractive for DNA computation, given that logic operations are constructed according to a simple binary combination of OR, NOT, and AND gates, for instance. This method allows higher-level calculations to be performed by combining a number of logic gates, with any logic gate capable of construction through combining multiple NAND (negative-AND) gates. In conventional DNA computation, the recognition of output molecules is mainly performed by combining several methodologies such as gel-electrophoretic or fluorescence detection following polymerase chain reaction (PCR) amplification.[1,3,4] To improve the speed of decoding, we have recently proposed nanopore decoding for detection of the output molecules directly and electrically. We constructed several logic gates including AND, OR, NOT, and NAND, and the output molecule was detected by nanopore measurement of the electrical signals in a droplet-based nanopore device.[5-7] In addition, we have also studied nanopore decoding for solving the Hamiltonian path problem with parallel computation, as mentioned above.[8] Through these studies, we are convinced that nanopore decoding is appropriate for rapid and simple decoding in DNA computing.

The field of DNA computing was developed largely as a curiosity-driven exercise focused on solving mathematics-related problems, including cryptograms and constructing various types of logic gates. Recently, this field increased in importance due to its potential applications in medical diagnosis.[9] Benenson *et al.* reported autonomous

diagnosis and drug-release systems with DNA computing using the following "if-then" logic: "if" certain diagnostic conditions are true, such as low expression levels of certain mRNAs relative to those of others, "then" the antisense drug is released.[10] After this pioneering study, several studies were undertaken focused on the application of this technology to diagnosis and therapy. Based on the favorable compatibility of nanopore technology with oligonucleotide detection, strategies utilizing this method for diagnosis using nanopores and DNA have been proposed.[11] MicroRNA, which is a short non-coding RNA that has about 18-25 nucleotides, is an important target in terms of diagnosis for cancers because its expression is regulated with cancer types or cancer stage, and the high-cancer specificity of the pattern of miRNA expression is attracting attention as a form of liquid biopsy.[12-14] In recent years, several approaches based on logic operations have been developed, including conventional DNA computation and gold nanoparticle strategies.[15,16] We have also constructed the AND gate for the detection of two overexpressed miRNAs (miR-20a and miR-17-5p) that are secreted from small cell lung cancer (SCLC).[17] In this system, two diagnostic DNAs were encapsulated in input droplets, and formed a four-way junction with the miRNAs only when the two miRNAs were present at the same time (equivalent to an AND gate operation). The structure of the four-way junction blocked the nanopore and generated a long current inhibition as the output signal, resulting in current blockings with diagnostic lengths of time for each pattern of (1, 1), (0, 1), (1, 0), (0, 0), as an AND gate. This logic operation is simple and it is useful in nanopore diagnosis. However, although parallel operation is the most intriguing characteristic in DNA computing, simple logic operation remains a challenge for multiplex diagnosis.

We here report a method for the identification of the expression patterns of five

different types of miRNAs (miR-193, miR-106a, miR-15a, miR-374, and miR-224) based on DNA computing combined with nanopore decoding.[18] These miRNAs are overexpressed in bile duct cancer (BDC), which is one of the highest mortality cancers. A diagnostic DNA with hairpin structure (HP-dgDNA) is employed as a computational molecule, and codes the expression pattern of the five miRNAs by forming a duplex structure with the miRNAs. The information encoded by HP-dgDNA is decoded by the unzipping of each miRNA complex in the nanopore and analysis of the ion current signals from the nanopore measurements. We analyzed the unzipping time of the miRNA pattern and were able to recognize the BDC-specific expression pattern of the miRNAs even in clinical samples. In addition, we found that our method can detect the miRNA pattern at the attomolar (10^{-18} M) level using an excess of HP-dgDNA.

Methods

Reagents and chemicals

All aqueous solutions were prepared with ultrapure water from a Milli-Q system (Millipore, Billerica, MA, USA). The reagents were as follows: 1, 2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC; Avanti Polar Lipids, Alabaster, AL, USA), *n*-decane (Wako Pure Chemical Industries, Ltd., Osaka, Japan), potassium chloride (KCl; Nacalai Tesque), 3-(N-Morpholino) propanesulfonic Acid (MOPS; Nacalai Tesque, Kyoto, Japan). Buffered electrolyte solutions (1.0 M KCl, 10 mM MOPS, pH 7.0) were prepared from ultrapure water. Wild-type alpha-hemolysin (αHL; List Biological Laboratories, Campbell, CA, USA, and Sigma-Aldrich, St Louis, MO, USA) was obtained as the monomer polypeptide, isolated from *Staphylococcus aureus* in the form of a powder and dissolved at a concentration of 1 mg/mL in ultrapure water. For use, samples were diluted

to the designated concentration using a buffered electrolyte solution and stored at 4°C. High-performance liquid chromatography (HPLC)-grade DNA oligonucleotides and miRNA were synthesized by FASMAC Co., Ltd. (Kanagawa, Japan) and stored at −20°C and −80°C, respectively.

Patients and healthy samples

Plasma samples were obtained from 6 patients with histologically proven BDC. All of the samples were obtained from patients who had undergone surgical resection in May 2010-November 2016. The diagnosis of these patients was based on histological assessment after surgical resection. Additionally, plasma samples were collected from 11 healthy volunteers (HVs) and a plasma pool was made of the 6 HVs. The clinicopathological backgrounds of the patients and HVs are shown in **Supplementary Table 1**. After sample collection, blood samples were centrifuged at 3,000 rpm for 10 min at 15°C to spin down the blood cells. Plasma samples were then transferred into fresh collection tubes and stored at -80°C until further processing.

Small RNA extraction

Small RNA was extracted from 300 µL of plasma with a NucleoSpin® miRNA Plasma (Takara) according to the manufacturer's protocol. At the beginning of each extraction procedure, exogenous control cel-miR-39-3p (FASMAC, Kanagawa, Japan) were spiked into samples before the addition of lysis buffer. The final volume was 30 µL. All eluted RNA samples were stored at -80°C until used.

Quantification of miRNA by quantitative real-time PCR

Amounts of miRNAs were quantified by quantitative real-time PCR (RT-qPCR) using the SYBR Advantage qPCR Premix (Takara Bio Inc., Shiga, Japan). The reverse transcription reaction was carried out with the Mir-X™ miRNA First-Strand Synthesis Kit (Takara)

according to the manufacturer's instructions. Quantitative PCR was performed on the Thermal Cycler Dice Real-Time System *Lite* (Takara), and reaction mixtures were incubated at 95℃ for 10 s, followed by 40 cycles at 95℃ for 5 s and 65℃ for 25 s. The cycle threshold (C_t) values were calculated with Multiplate RQ (Takara).

HP-dgDNA/miRNA hybridization

Diagnostic solutions consisted of each extracted miRNA with 500 nM HP-dgDNA in MOPS buffer (pH 7.0, 10 mM) containing 1 M potassium chloride. These solutions were heated to 95°C for 5 min and then cooled to room temperature gradually.

Preparation of the microdevice

The microdevice was fabricated by machining a 6.0 mm thick, 10×10 mm polymethyl methacrylate (PMMA) plate (Mitsubishi Rayon, Tokyo, Japan) using a computer-aided design and computer-aided manufacturing-three-dimensional modeling machine (MM-100, Modia Systems, Japan). Two wells (2.0 mm diameter and 4.5 mm depth) and a chase between the wells were manufactured on the PMMA plate. Each well had a through-hole in the bottom and Ag/AgCl electrodes were set into this hole (**Figure 1a**). A polymeric film made of parylene C (polychloro-*p*-xylylene) with a thickness of 5 μm was patterned with single pores (100 μm diameter.) using a conventional photolithography method and then fixed between PMMA films (0.2 mm thick) using an adhesive bond (Super X, Cemedine Co., Ltd, Tokyo, Japan). The films, including the parylene film, were inserted into the chase to separate the wells. Pico2 (Tecella, Foothill Ranch, CA, USA) was used for current measurements.

Bilayer lipid membrane preparation and reconstitution of αHL

Bilayer lipid membranes (BLMs) were prepared using a device produced by microfabrication (**Figure 1a**). BLMs can be simultaneously formed in this device by the

droplet contact method (**Figure 1b**). [19-24] In this method, the two lipid monolayers contact each other and form BLMs on a parylene C film that separates two chambers. BLMs were formed as follows: the wells of the device were filled with n-decane (2.5 µL) containing DPhPC (10 mg/mL). The recording solutions (4.7 µL) on each side of the BLMs contained 1 M KCl and 10 mM MOPS (pH 7.0). αHL was reconstituted in BLMs to form a nanopore from the ground side. The diagnostic solution after HP-dgDNA/miRNAs hybridization was also added to the ground side. Within a few minutes of adding the solutions, BLMs were formed and αHL created nanopores within them. When the BLMs ruptured during this process, they were recreated by tracing with a hydrophobic stick at the interface of the droplets.

Channel current measurements and data analysis

The channel current was recorded with an Axopatch 200B amplifier (Molecular Devices, USA), filtered with a low-pass Bessel filter at 10 kHz with a sampling rate of 50 kHz. A constant voltage of +200 mV was applied from the recording side, and the ground side was grounded. The recorded data from Axopatch 200B were acquired with Clampex 9.0 software (Molecular Devices, USA) through a Digidata 1440A analog-to-digital converter (Molecular Devices, USA). Data were analyzed using Clampfit 10.6 (Molecular Devices, USA), Excel (Microsoft, Washington, USA), and Origin Pro 8.5J (Light Stone, Tokyo, Japan). DNA or miRNA translocation and blocking were detected when >80% of open αHL channel currents were inhibited. Between 251 and 322 translocating or blocking events were recorded. From this data, we generated the histograms of unzipping time for each sample using a bootstrap method. The event frequency was counted for each 1-min interval when one pore was open. Nanopore measurements were conducted at $22 \pm 2°C$.

Stability prediction of the miRNA/HP-dgDNA duplex

The thermodynamic stability of the miRNA/HP-dgDNA duplex was predicted by the nearest-neighbor (NN) model with NN parameters for RNA/DNA hybrids.[25] The summation of the free energy changes due to helix initiation and for each subsequent base pair in the sequence were calculated as the Gibbs free energy (ΔG_{sim}) of the duplex.

Results

Design of diagnostic DNA (HP-dgDNA)

We selected miR-193, miR-106a, miR-15a, miR-374, and miR-224 as target miRNAs for BDC diagnosis because these miRNAs have been reported to be overexpressed in human intrahepatic cholangiocarcinoma.[18] We designed HP-dgDNA, a diagnostic DNA, with a hairpin structure and with a sequence such that the five miRNAs can linearly bind to the diagnostic DNA (**Figure 1c, Supplementary Table 2**). The detailed design processes are as follows:

(i) The complementary strands of the 5 miRNAs are inserted into the main sequence to encode the miRNA patterns (dgDNA).

(ii) Poly(dC)$_{20}$ is added at the 3' end of dgDNA to be exclusively inserted into αHL pore. The length of poly(dC)$_{20}$ is ca. 8.4 nm; therefore, the HP-dgDNA can penetrate to the αHL pore because the length from the entrance to the β-barrel structure of the pore is 4.8 nm.

(iii) The HP-dgDNA hairpin structure is added at the 5' end of dgDNA to prevent insertion into the αHL pore from the 5' end.[26]

All designed structures were checked by thermodynamic simulation (NUPACK: http://www.nupack.org/, and Nearest Neighbor (NN) model). The duplex formation of

8

HP-dgDNA and the miRNAs at 500 nM each were simulated thermodynamically (**Figure 1d**). The NUPACK simulation can calculate ΔG_{sim} of a secondary structure composed of only DNA or RNA alone. Therefore, ΔG_{sim} of DNA/RNA binding was calculated by the NN model using NN parameters.[25] All ΔG_{sim} of each miRNA/HP-dgDNA hybridizations are listed in **Supplementary Table 3**.

Two cytosines were set in-between the complimentary sequence of each miRNAs as spacers. The thermodynamic simulation showed that the two-cytosine spacers make the duplex sufficiently stabilized (**Supplementary Figure 1**). We next determined the order of complementary strands of the 5 miRNAs in HP-dgDNA using simulations. To inhibit the unpredicted formation of secondary structures, the sequence order was selected to have the smallest hybridization energy from the 120 (=5!) possible sequence orders; the smallest ΔG_{sim} was −94.8 kJ mol^{-1} in the optimal order (**Supplementary Figure 2**). The HP-dgDNA structure fulfilling the above requirements is shown in **Figure 1c** and **1d**.

The long cytosine homo-sequence is selected as the 3'-end-tail because we can detect the insertion of HP-dgDNA from the unique current blocking ratio (I_b=85%) of poly(dC).[27] The blocking ratio of poly(dC) is likely to be higher than 70% at an unobstructed level, while the blocking ratio by poly(dA) and poly(dT) is likely to be 50% of the unobstructed level. Therefore, we planned to analyze ion current signals inhibited over 80% and detect the target molecule of the HP-dgDNA/miRNAs duplex (**Figure 1e**).
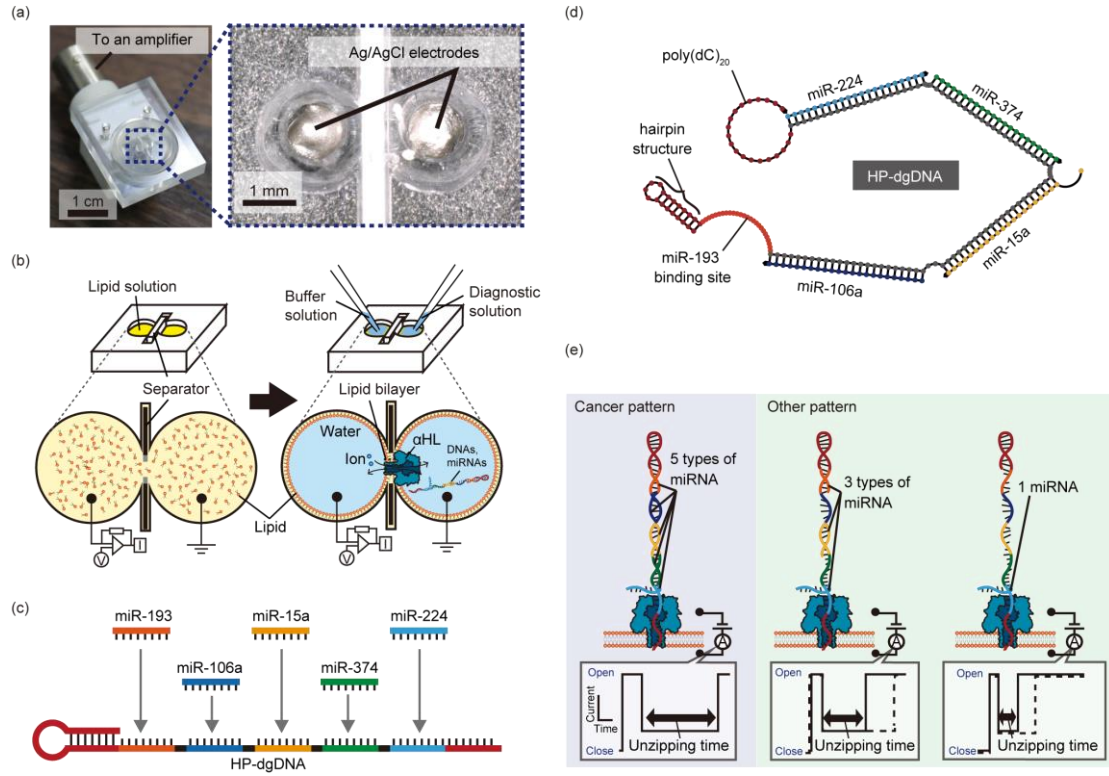
Figure 1 (a) Photographs of the devices for nanopore measurement. (b) Schematic illustration of lipid bilayer preparation. (c) Design of HP-dgDNA. (d) Structure of the duplex of HP-dgDNA and miR-193, miR-106a, miR-15a, and miR-374, simulated by NUPACK. When we simulated the structure, we converted the sequence of HP-dgDNA into the corresponding RNA sequence and replaced the 'A' of HP-dgDNA sequence with 'U'. (e) Schematic illustration of nanopore decoding.

4.2 Nanopore measurement of each miRNA pattern

In the nanopore measurement of the HP-dgDNA/miRNA duplex, there were two possibilities: that the duplex passed through the nanopore with unzipping, or that the duplex returned to the *cis* solution.[28] In order to confirm translocation of HP-dgDNA of the duplex, we measured the complex of HP-dgDNA/miR-193 using αHL at 150 mV or 200 mV (**Supplementary Figure 3a**). Histograms of the duration time were made using bootstrapping, a statistical method, and the mean value of the duration time at 200 mV was shorter than that at 150 mV (**Supplementary Figure 3b**). The duration time was shortened by increasing the applied voltage, suggesting that the duplex of HP-dgDNA/miR-193 entered the pore, was pushed out by the voltage, and then passed through the pore with unzipping of miR-193. The unzipping signal ratio has been reported to become higher with an increase in applied voltage.[29,30] Therefore, in order to facilitate the unzipping, the applied voltage of 200 mV was adopted for the following experiments.

To validate the proof of concept, we prepared a cancer miRNA pattern (with all 5 miRNAs present) and two healthy miRNA patterns (with 1 or 3 miRNAs present) using synthetic miRNAs. In each pattern, unzipping signals with over 80% inhibition were observed, hence we propose that HP-dgDNA was inserted from poly(dC)$_{20}$ and subsequently passed through the αHL pore with unzipping of the miRNAs (**Figure 2a-2c**). We analyzed the unzipping time with the consideration that unzipping time represents the number of miRNAs bound to HP-dgDNA. Histograms of the unzipping time were also made using bootstrapping (**Figure 2d**). The peak value of unzipping-time histogram of miR-374, miR-15a, miR-224, miR-106a, miR-193, three miRNAs, and five miRNAs was 709 ms, 736 ms, 1081 ms, 1349 ms, 2052 ms, 4517 ms, and 5841 ms, respectively.

Each peak value was larger than that of HP-dgDNA itself, at 369 ms. This result shows that miRNAs were unzipped from HP-dgDNAs translocating the αHL pore, each with a characteristic unzipping time after binding to HP-dgDNAs (**Figure 2e**). The unzipping time became larger as the number of miRNAs binding to HP-dgDNA increased, and corresponded to ΔG_{sim} of the duplex. In the case of the singly-bound miRNA, the unzipping time also became longer with increasing ΔG_{sim} (**Figure 2e, 2f**). These results indicate that complexes with a larger ΔG_{sim} require a longer time to be unzipped from HP-dgDNA in the αHL pore. The unzipping time showed a logarithmic dependence on ΔG_{sim} ($R^2 = 0.82$), and this result suggests that our system could be used to recognize other miRNA patterns of different cancers by adjusting ΔG_{sim} and controlling the unzipping time (**Figure 2g**).
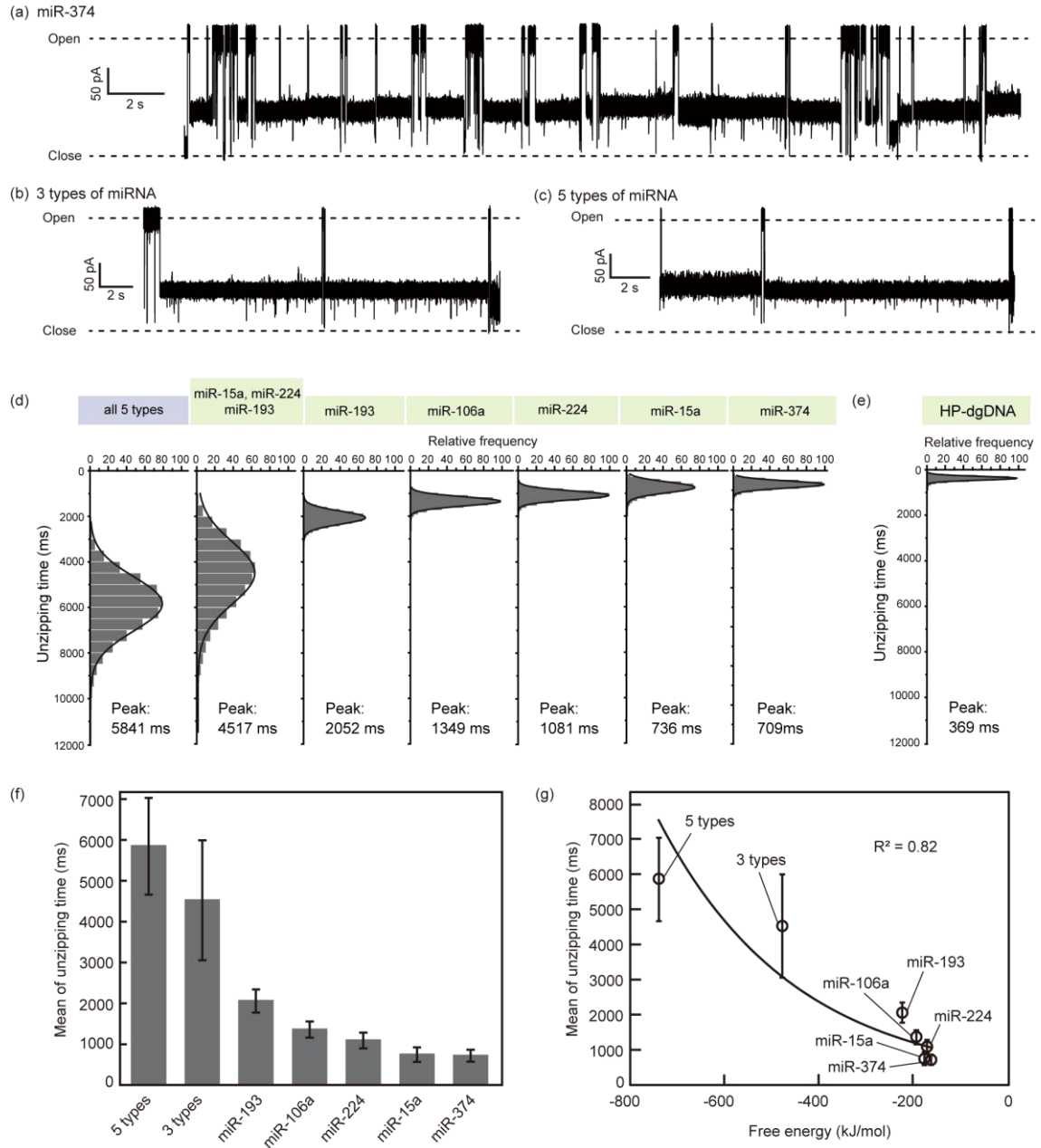
Figure 2   (a) Characteristic current signal of miR-374. (b) Characteristic current signal of the pattern of miR-15a, miR-224, and miR-193. (c) Characteristic current signal of the pattern of miR-15a, miR-224, miR-193, miR-106a, and miR-374. (d) Histograms of unzipping time of each miRNA pattern. (e) Histogram of duration time of HP-dgDNA. (f) Mean of the unzipping time of each miRNA pattern. (g) Mean of the unzipping time as a function of the free energy of each duplex of HP-dgDNA and miRNAs. Error bars are mean ± SD after bootstrapping.

13

4.3 RT-qPCR of healthy and cancer samples

Based on the results of the proof of concept study, we next attempted to recognize miRNA patterns using a clinical plasma sample. The five miRNAs, miR-15a, miR-193, miR-224, miR-106a, miR-374 are known to increase in cholangiocytes microdissected from the tissue of BDC patients.[18] To confirm the expression level in plasma before performing the nanopore decoding, absolute quantification of miRNAs extracted from the plasma samples was conducted by RT-qPCR. First, we prepared calibration curves of the five miRNAs and cel-miR-39-3p, as the spike-in miRNA (**Supplementary Figure 4**). The difference in efficiency of reverse transcription and amplification between the five miRNAs was normalized using the calibration curves of the five miRNAs. The difference in efficiency of reverse transcription and amplification between the samples was normalized by the calibration curve of the spike-in miRNA. The miRNA concentration in each of the samples was between approximately 30 aM to 7.0 nM (**Figure 3a**). To distinguish the cancer patients and the healthy volunteers (HVs), we had to interpret the five dimension data (= concentration value of the five miRNAs). One way to distinguish cancer patients from healthy controls using a combination of values attributed to multiple miRNAs is by logistic regression.[31] We also considered it desirable to use logistic regression for our analysis, but it has been reported that the sample size for logistic regression should be at least five times larger than the explanatory variables.[32] For this study, we therefore needed more than $5 \times 5 = 25$ samples because there were five types of miRNAs. Since clinical samples of BDC are difficult to obtain due to the small patient numbers in Japan, we instead averaged the concentration value of the miRNAs of the five cancer patients and the five HVs for analysis. The concentration of each of the miRNAs

in the cancer patients was higher than in the HVs (**Figure 3b**). We thus revealed that the

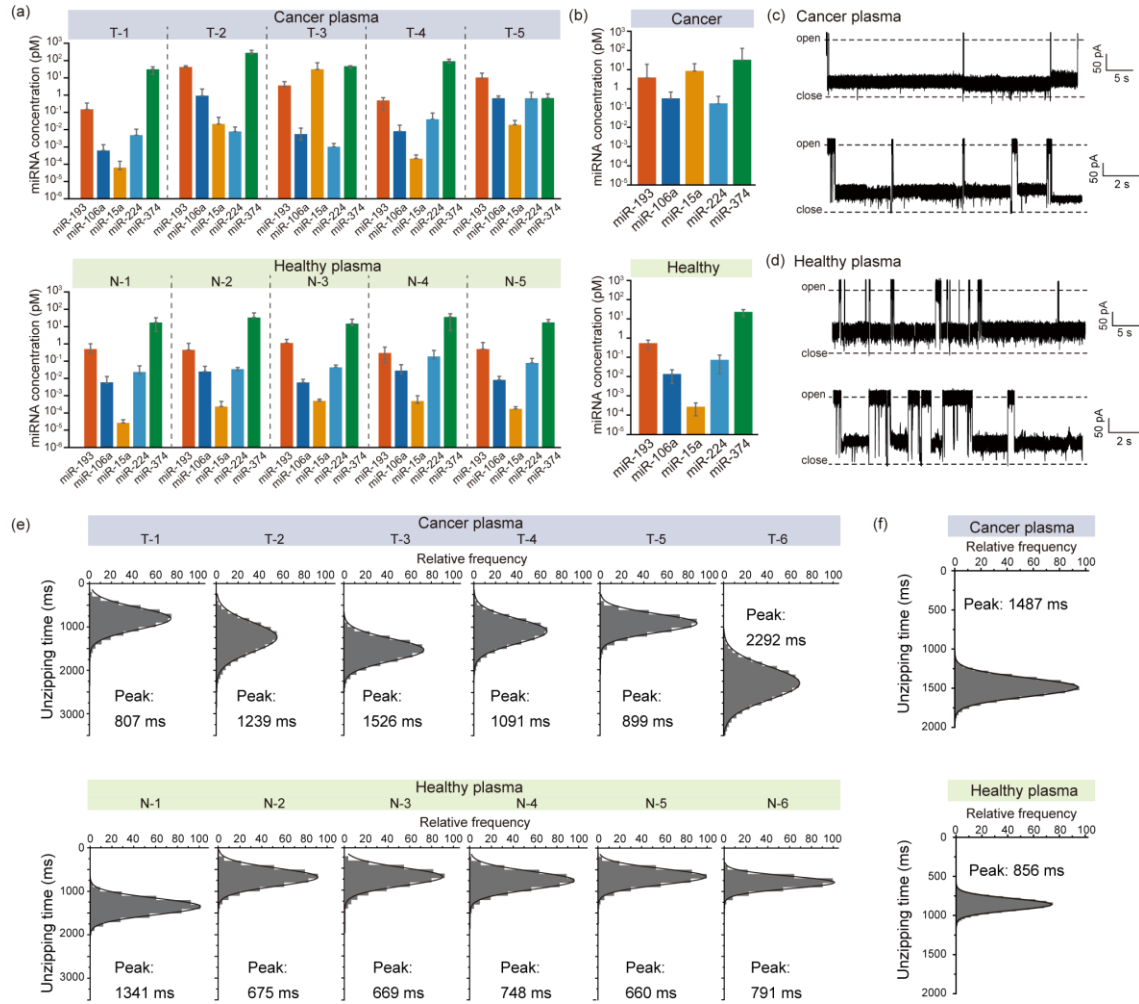five miRNAs known to increase in BDC tissue also overexpressed in plasma.

Figure 3 (a) miRNA concentration of the 5 cancer plasma samples and the 5 healthy plasma samples quantified by RT-qPCR. (b) The average miRNA concentration of the 5 cancer samples and the 5 healthy samples. Characteristic current signal of cancer plasma samples (c) and healthy plasma samples (d). (e) Histograms of the unzipping time of the 6 cancer samples and the 6 healthy samples. (f) The average unzipping time of plasma samples. (e) miRNA concentration of the 5 cancer plasma samples and the 5 healthy plasma samples quantified by RT-qPCR. (f) The average miRNA concentration of the 5 cancer samples and the 5 healthy samples. All error bars are mean ± SE.

4.4 Nanopore measurement of healthy and cancer samples

HP-dgDNA hybridized with miRNAs extracted from the plasma samples were examined by nanopore measurement. In both cancer and HV samples, over 80% inhibition of the ion current flowing through the pore was observed, as also observed with synthetic miRNAs (**Figure 3c-3d**), indicating that miRNA detection is possible even with use of clinical samples. It is possible that HP-dgDNA and the five miRNAs can compose 120 (= 5!) types of secondary structure in the hybridized combinations with different concentrations. We observed various unzipping times on the order of 10^{-2} to 10^{6} ms (**Supplementary Figure 5**), probably due to the wide variety of hybridized structures. We assumed that distributions of unzipping time represented the overall miRNA pattern in each sample, since we consider the ratio of HP-dgDNA/miRNAs duplex in a crowd of HP-dgDNA to be reflected in the distributions. Therefore, the histogram of the unzipping time was prepared by bootstrapping and the peak value of each of the cancer patients and HVs were obtained (**Figure 3e**). The average unzipping time of the cancer patients and HVs were 1487 ms and 856 ms, respectively (**Figure 3f**). The miRNA expression level was higher in the cancer patients than the HVs as in the result of RT-qPCR. This result indicates that our system could recognize the characteristic miRNA pattern of bile duct cancer.

Surprisingly, our nanopore system could discriminate the clinical samples which included the low-concentration (aM) miRNAs, whose concentration was confirmed by RT-qPCR as previously mentioned. This result came as a surprise as conventional nanopore measurements have been considered to have limited ability to detect nucleic acids at concentrations lower than 0.1 pM.[28] However, in this study, our nanopore

measurements seem to detect sub-femtomolar miRNAs directly. Next, we attempted to confirm the detection of sub-femtomolar concentration using synthetic miRNAs *in vitro*.
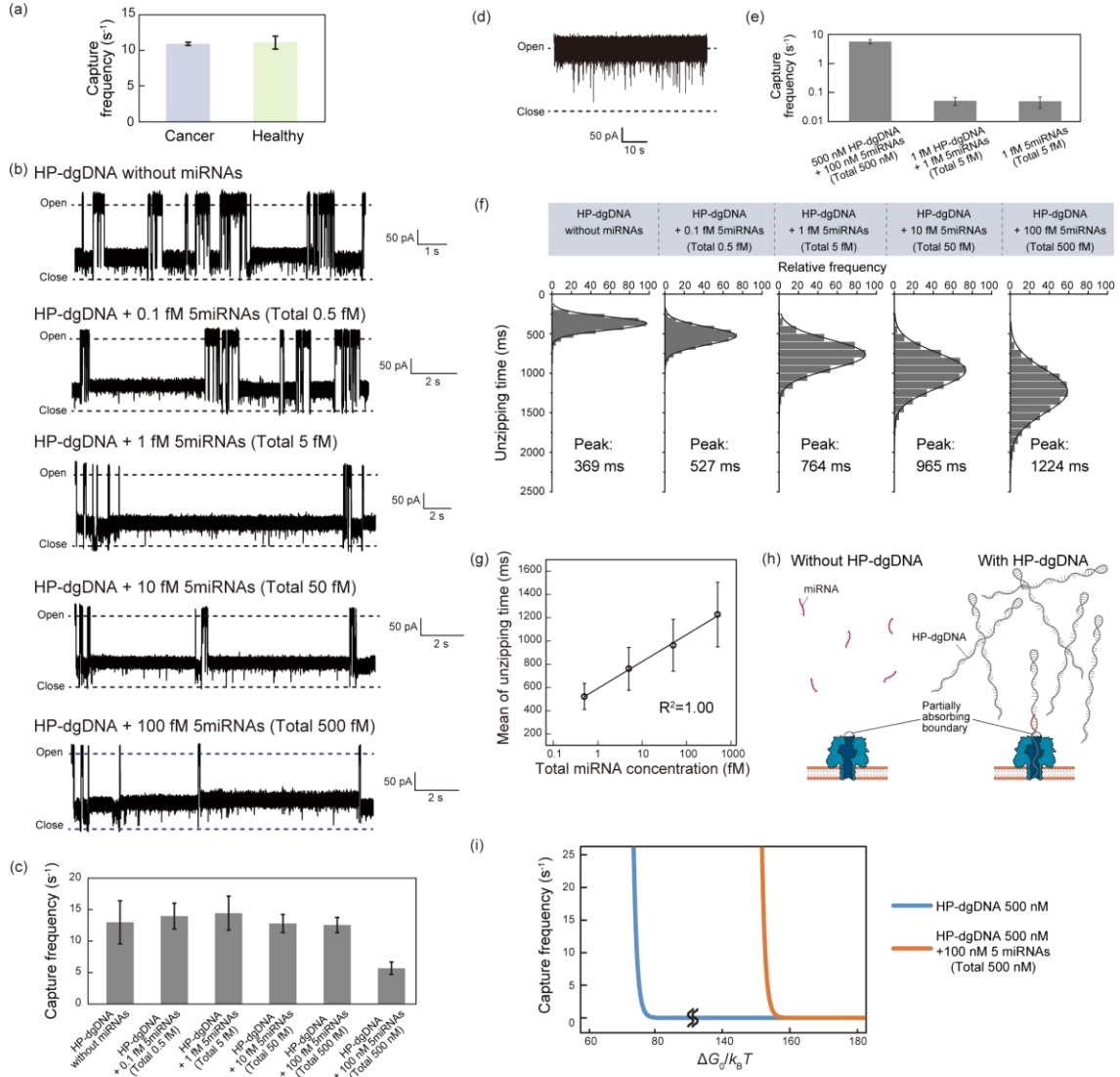

Figure 4 (a) The average capture frequency of the 6 cancer samples and the 6 healthy samples in the presence of dgDNA. (b) Characteristic current signals of 500 nM HP-dgDNA and miRNAs at each concentration. (c) Capture frequency of 500 nM HP-dgDNA and miRNAs for each condition. Using 5 fM miRNAs without HP-dgDNA, we observed a characteristic signal (d) and its capture frequency (e). (f) Histograms of the unzipping time of each miRNA pattern. HP-dgDNA was 500 nM in each. (g) Mean of the unzipping time as a function of the total miRNA concentrations. All error bars are mean ± SD. (h) Schematic illustration of the femtomolar miRNA detection with/without 500 nM HP-dgDNA. (i) The capture frequency calculated by the theoretical model.

4.5 Investigation of sub-femtomolar detection with a model system using synthetic miRNAs

Generally, the capture frequency of DNA/RNA translocation *via* a nanopore is used to estimate the detection capability against a target concentration. We also calculated the capture frequency of the plasma samples in addition to the unzipping times. However, as shown in **Figure 4a**, monitoring the capture frequency does not allow for recognition of the miRNA patterns in the BDC, because the HP-dgDNA is present at higher concentration (500 nM) and therefore dominates the frequency events. To investigate detection at low concentration, we designed model experiments with similar experimental conditions to the clinical experiments: 0.1, 1, 10, and 100 fM of each of the 5 miRNAs were used with 500 nM of HP-dgDNA and the capture frequency was analyzed (**Figure 4b**). Although the capture frequencies of the femtomolar conditions were similar to each other, the 500 nM HP-dgDNA/miRNAs duplex showed a lower frequency (**Figure 4c**). These results suggest that miRNAs/HP-dgDNA with residual HP-dgDNA show higher frequency and miRNAs/HP-dgDNA itself shows lower frequency. Besides, in the presence of only miRNAs without HP-dgDNA, the frequency of 1 fM 5miRNAs (total 5 fM) with over 80% inhibition was around 0.05 s^{-1}; this value was less than one-hundredth of 500 nM HP-dgDNA/5miRNAs (5.68 s^{-1}, **Figure 4d, 4e**).

We next analyzed the unzipping time of the model experiments. The unzipping times of the femtomolar concentrations of miRNAs with 500 nM HP-dgDNA were measured and compared with each other. The unzipping time increased with the increase in concentration of the miRNAs, suggesting that 0.1 fM miRNA can be detected using the HP-dgDNA at sub-femtomolar concentrations (**Figure 4f, 4g**).

5. Discussion

Low concentration detection of miRNA/HP-dgDNA in the nanopore measurements

Our results when using clinical samples were a pleasant surprise, since nanopore measurements have generally been considered to be unable to detect nucleic acids at lower than picomolar concentration.[28] To measure the sub-picomolar or femtomolar targets, several elaborated methods has been proposed; Wang *et al*. showed detection of 0.1 pM of miR-155 with the αHL nanopore using asymmetric solution conditions of 0.2 M/3 M (*cis/trans*) KCl salt gradient, to increase the capture frequency of the target molecule.[28] Zhang *et al*. used the αHL nanopore with isothermal amplification of nucleic acids, resulting in the detection of 1 fM of miR-20a.[33] In this study, we did not use such methodologies, neither the salt gradient nor the amplification of nucleic acids, but our method seems to detect sub-femtomolar miRNAs directly.

We consider that the key to this low concentration detection is the excess amount of HP-dgDNA as the complementary probe for the target miRNAs. Our results in the *in vitro* experiments support this hypothesis:

1) In the case of the equimolar condition between the HP-dgDNA and miRNAs or the miRNAs themselves, the total capture frequency is constant or decreased (**Figure 4c**).

2) In the case of the excess amount of the HP-dgDNA, the unzipping time showed a linear relationship for concentrations ranging from 0.1 fM to 1 pM (**Figure 4g**).

Consequently, excess HP-dgDNA surrounding target HP-dgDNA/miRNAs would cause an increase in the sensitivity, as schematically shown in **Figure 4h**. To support this hypothesis from a theoretical view, we refer to a nanopore capture model as we have recently proposed.[34] The modeling of the capture frequency *f* of a particle captured into a nanopore is described by the following equation,

$$\frac{1}{f} = \frac{1}{f_a} + \frac{1}{f_e} \quad , \tag{1}$$

with an approach frequency f_a related to the migration of particles from the bulk to the pore inlet, and an entrance frequency f_e related to the actual entry of particles into the pore region. The f_a and f_e are defined as follows:

$$f_a = \frac{2\pi a_2 C_0}{1 - e^{-\frac{a_2}{r_e D}}} \quad , \tag{2}$$

with

$$a_2 = \frac{\mu q I}{2\pi\sigma} + \frac{Q_f}{2\pi} \quad , \tag{3}$$

where C_0 is the particle concentration, r_e is the pore entrance radius, and D denotes the diffusion coefficient, μ is the particle mobility, q is the particle charge, I is the ion current flowing through the pore, σ is the electrolyte conductivity and Q_f is the volumetric flow rate entering the pore, and

$$f_e = 2\pi C_0 r_e^3 \frac{k_B T}{h} \exp\left[2\varphi(r_e) - \frac{\Delta G_0}{k_B T}\right] \quad , \tag{4}$$

where k_B is the Boltzmann constant, T is the temperature, h is the Planck constant, $\varphi(r_e)$ is the dimensionless effective potential evaluated at r_e, and ΔG_0 the free energy barrier at equilibrium. When calculating f for HP-dgDNA/5miRNAs and HP-dgDNA respectively, the parameters that differ between them are q and ΔG_0. If we first consider q: if the ratio of the HP-dgDNA/5miRNAs signal was 10 out of 300 signals for the experiments using 0.5 fM miRNAs, f for HP-dgDNA/5miRNAs would become more than 10^6 times larger than that for HP-dgDNA. However, using the elemental charge e in this assumption, we calculated that $q = 270e$ for HP-dgDNA/5miRNAs and $q = 160e$ for HP-dgDNA; indicating that q is unlikely to have such a significant effect on f. Therefore, we assumed that the free energy barrier at equilibrium ΔG_0 between HP-dgDNA/5miRNAs and HP-

dgDNA is considerably different and affects f. Although proper modeling to predict ΔG_0 is required, ΔG_0 can be represented as:

$$\Delta G_0 = c \times k_{\mathrm{B}}T \tag{5}$$

with c a target-specific constant. Using this model for our case, we plotted f against $\Delta G_0/k_{\mathrm{B}}T \, (= c)$ in **Figure. 4i.** As a result, c of ssDNA and of dsDNA captured into a nanopore is expected to be significantly different, suggesting that ΔG_0 differs between ssDNA and dsDNA (**Figure 4i**). Therefore, a large excess of ssDNA could promote the migration of dsDNA to the pore.

6. Conclusion

In conclusion, we propose a system for miRNA pattern recognition using DNA computing and nanopore decoding. The information encoded in miRNA hybridized to HP-dgDNA is decoded by nanopore sensing in the form of unzipping time. We successfully distinguished miRNA expression patterns in the clinical plasma samples of bile duct cancer patients and healthy volunteers. In addition, we found that our system could detect 0.1 fM miRNAs using HP-dgDNA of a higher concentration than the miRNAs. We believe that our system could improve the sensitivity of nanopore sensing and be a simple diagnostic tool for cancer.

7. Acknowledgements

8. References

1   Adleman, L. M. Molecular computation of solutions to combinatorial problems. *Science* **266**, 1021-1024, doi:10.1126/science.7973651 (1994).

2   Xu, X. H., Shang, Y. X., Liu, F. S., Jiang, Q. & Ding, B. Q. Logic devices based on nucleic acid self-assembly. *Infomat* **3**, 1070-1082, doi:10.1002/inf2.12240 (2021).

3   Benenson, Y. *et al.* Programmable and autonomous computing machine made of biomolecules. *Nature* **414**, 430-434, doi:10.1038/35106533 (2001).

4   Chen, Y. Q. *et al.* A DNA logic gate based on strand displacement reaction and rolling circle amplification, responding to multiple low-abundance DNA fragment input signals, and its application in detecting miRNAs. *Chem Commun* **51**, 6980-6983, doi:10.1039/c5cc01389e (2015).

5   Yasuga, H. *et al.* Logic Gate Operation by DNA Translocation through Biological Nanopores. *Plos One* **11**, e0149667, doi:10.1371/journal.pone.0149667 (2016).

6   Ohara, M., Takinoue, M. & Kawano, R. Nanopore Logic Operation with DNA to RNA Transcription in a Droplet System. *Acs Synthetic Biology* **6**, 1427-1432, doi:10.1021/acssynbio.7b00101 (2017).

7   Hiratani, M. & Kawano, R. DNA logic operation with nanopore decoding to recognize microrna patterns in small cell lung cancer. *Analytical Chemistry* **90**, 8531-8537, doi:10.1021/acs.analchem.8b01586 (2018).

8   Takiguchi, S. & Kawano, R. Nanopore decoding for a Hamiltonian path problem. *Nanoscale* **13**, 6192-6200, doi:10.1039/D0NR09031J (2021).

9   Benenson, Y. Biomolecular computing systems: principles, progress and potential. *Nat. Rev. Genet.* **13**, 455-468, doi:10.1038/nrg3197 (2012).

10   Benenson, Y., Gil, B., Ben-Dor, U., Adar, R. & Shapiro, E. An autonomous molecular computer for logical control of gene expression. *Nature* **429**, 423-429, doi:10.1038/nature02551 (2004).

11   Kawano, R. Nanopore Decoding of Oligonucleotides in DNA Computing. *Biotechnology Journal* **13**, 1800091, doi:10.1002/biot.201800091 (2018).

12   Cai, S. *et al.* Single-molecule amplification-free multiplexed detection of circulating microRNA cancer biomarkers from serum. *Nat Commun* **12**, 3515, doi:10.1038/s41467-021-23497-y (2021).

13   Liyanage, T. *et al.* Plasmoelectronic-Based Ultrasensitive Assay of Tumor Suppressor microRNAs Directly in Patient Plasma: Design of Highly Specific Early Cancer Diagnostic Technology. *Analytical Chemistry* **91**, 1894-1903, doi:10.1021/acs.analchem.8b03768 (2019).

14   Wang, H., Peng, R., Wang, J. J., Qin, Z. L. & Xue, L. X. Circulating microRNAs as potential cancer biomarkers: the advantage and disadvantage. *Clinical Epigenetics* **10**, doi:10.1186/s13148-018-0492-1 (2018).

15   Zhang, C. *et al.* Cancer diagnosis with DNA molecular computation. *Nature Nanotechnology* **15**, 709-715, doi:10.1038/s41565-020-0699-0 (2020).

16   Liu, L. *et al.* Gold Nanoflares with Computing Function as Smart Diagnostic Automata for Multi-miRNA Patterns in Living Cells. *Analytical chemistry* **92**, 10925-10929, doi:10.1021/acs.analchem.0c02325 (2020).

17   Hiratani, M., Ohara, M. & Kawano, R. Amplification and Quantification of an Antisense Oligonucleotide from Target microRNA Using Programmable DNA and a Biological Nanopore. *Analytical chemistry* **89**, 2312-2317, doi:10.1021/acs.analchem.6b03830 (2017).

18   Chen, L. *et al.* The role of microRNA expression pattern in human intrahepatic cholangiocarcinoma. *J. Hepatol.* **50**, 358-369, doi:10.1016/j.jhep.2008.09.015 (2009).

19   Watanabe, H. & Kawano, R. Channel Current Analysis for Pore-forming Properties of an Antimicrobial Peptide, Magainin 1, Using the Droplet Contact Method. *Anal Sci* **32**, 57-60, doi:10.2116/analsci.32.57 (2016).

20   Kawano, R. *et al.* A portable lipid bilayer system for environmental sensing with a transmembrane protein. *PLoS One* **9**, e102427, doi:10.1371/journal.pone.0102427 (2014).

21   Tsuji, Y. *et al.* Droplet split-and-contact method for high-throughput transmembrane electrical recording. *Analytical chemistry* **85**, 10913-10919, doi:10.1021/ac402299z (2013).

22   Tsuji, Y. *et al.* Droplet-based lipid bilayer system integrated with microfluidic channels for solution exchange. *Lab on a chip* **13**, 1476-1481, doi:10.1039/c3lc41359d (2013).

23   Kawano, R. *et al.* Automated Parallel Recordings of Topologically Identified Single Ion Channels. *Scientific Reports* **3**, 1995, doi:10.1038/srep01995 (2013).

24   Kawano, R., Osaki, T., Sasaki, H. & Takeuchi, S. A Polymer-Based Nanopore-Integrated Microfluidic Device for Generating Stable Bilayer Lipid Membranes. *Small* **6**, 2100-2104, doi:10.1002/smll.201000997 (2010).

25   Banerjee, D., Tateishi-Karimata, H., Ohyama, T., Ghosh, S., Endoh, T., Takahashi, S., Sugimoto, N. Improved nearest-neighbor parameters for the stability of RNA/DNA hybrids under a physiological condition. *Nucleic Acids Research* **48**, 12042-12054, doi:https://doi.org/10.1093/nar/gkaa572 (2020).

26   Vercoutere, W. *et al.* Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nature Biotechnology* **19**, 248-252, doi:10.1038/85696 (2001).

27   Butler, T. Z., Gundlach, J. H. & Troll, M. Ionic current blockades from DNA and RNA molecules in the alpha-hemolysin nanopore. *Biophys J* **93**, 3229-3240, doi:10.1529/biophysj.107.107003 (2007).

28   Wang, Y., Zheng, D. L., Tan, Q. L., Wang, M. X. & Gu, L. Q. Nanopore-based detection of circulating microRNAs in lung cancer patients. *Nature Nanotechnology* **6**, 668-674, doi:10.1038/nnano.2011.147 (2011).

29   Mathe, J., Visram, H., Viasnoff, V., Rabin, Y. & Meller, A. Nanopore unzipping of individual DNA hairpin molecules. *Biophys J* **87**, 3205-3212, doi:10.1529/biophysj.104.047274 (2004).

30   Wang, Y., Tian, K., Hunter, L. L., Ritzo, B. & Gu, L. Q. Probing molecular pathways for DNA orientational trapping, unzipping and translocation in nanopores by using a tunable overhang sensor. *Nanoscale* **6**, 11372-11379, doi:10.1039/c4nr03195d (2014).

31   Yokoi, A. *et al.* A combination of circulating miRNAs for the early detection of ovarian cancer. *Oncotarget* **8**, 89811-89823, doi:10.18632/oncotarget.20688 (2017).

32   Vittinghoff, E. & McCulloch, C. E. Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology* **165**, 710-718, doi:10.1093/aje/kwk052 (2007).

33   Zhang, H., Hiratani, M., Nagaoka, K. & Kawano, R. MicroRNA detection at femtomolar concentrations with isothermal amplification and a biological nanopore. *Nanoscale* **9**, 16124-16127, doi:10.1039/c7nr04215a (2017).

34   Chinappi, M., Yamaji, M., Kawano, R. & Cecconi, F. Analytical Model for Particle Capture in Nanopores Elucidates Competition among Electrophoresis, Electroosmosis, and Dielectrophoresis. *Acs Nano* **14**, 15816-15828, doi:10.1021/acsnano.0c06981 (2020).