Direct microRNA sequencing using Nanopore Induced Phase-Shift Sequencing (NIPSS)

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

PII: S2589-0042(20)30100-0
DOI: https://doi.org/10.1016/j.isci.2020.100916
Reference: ISCI 100916

To appear in: ISCIENCE

Received Date: 15 January 2020
Revised Date: 31 January 2020
Accepted Date: 11 February 2020

Please cite this article as: Zhang, J., Yan, S., Chang, L., Guo, W., Wang, Y., Wang, Y., Zhang, P., Chen, H.-Y., Huang, S., Direct microRNA sequencing using Nanopore Induced Phase-Shift Sequencing (NIPSS), ISCIENCE (2020), doi: https://doi.org/10.1016/j.isci.2020.100916.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 The Author(s).
Direct microRNA sequencing using Nanopore Induced Phase-Shift Sequencing (NIPSS)

Jinyue Zhang,[1, 2], + Shuanghong Yan,[1, 2], + Le Chang,[1, 2] Weiming Guo,[1, 2, 3] Yuqin Wang,[1, 2] Yu Wang,[1, 2] Panke Zhang,[1, 2] Hong-Yuan Chen,[1] Shuo Huang[1, 2, 4], *

[1] State Key Laboratory of Analytical Chemistry for Life Sciences, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China.
[2] Chemistry and Biomedicine Innovation Center (ChemBIC), Nanjing University, Nanjing 210023, China.
[3] Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Department of Chemical Biology, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.
[4] Lead Contact

*Correspondence: shuo.huang@nju.edu.cn
+These authors contributed equally to this work

Keywords
MiRNAs, IsomiRs, Sequencing, Nanopore, N6-methyladenosine
SUMMARY

MicroRNAs (miRNAs) are a class of short non-coding RNAs that function in RNA silencing and post-transcriptional gene regulation. However, direct characterization of miRNA is challenging due to its unique properties such as its low abundance, sequence similarities and short length. Though urgently needed, single molecule sequencing of miRNA has never been demonstrated, to the best of our knowledge. Nanopore Induced Phase Shift Sequencing (NIPSS), which is a variant form of nanopore sequencing, could directly sequence any short analytes including miRNA. In practice, NIPSS clearly discriminates between different identities, isoforms and epigenetic variants of model miRNA sequences. This work thus demonstrates direct sequencing of miRNA, which serves as a complement to existing miRNA sensing routines by the introduction of the single molecule resolution. Future engineering of this technique may assist miRNA based early stage diagnosis or inspire novel cancer therapeutics.
INTRODUCTION

MicroRNAs (miRNAs) are a group of short, single-stranded, non-coding RNAs that act as post-transcriptional gene regulators for a wide variety of physiological processes, including proliferation, differentiation, apoptosis, and immune reactions (Kim et al., 2018; Mehta and Baltimore, 2016). On the other hand, aberrant miRNA expression levels have been shown to be closely related to diverse diseases, such as cancer (Bracken et al., 2016; He et al., 2015; Nassar et al., 2017; Teng et al., 2017), auto-immune disorders (Pauley et al., 2009) and inflammatory diseases (Hatziapostolou et al., 2011).

MiRNAs function by binding to the 3’ untranslated region (3’UTR) of target messenger RNAs (Dong et al., 2013). Minor sequence variations, which include trimming, addition or substitution of miRNA sequences, alter its binding affinities to target messenger RNA (Dong et al., 2013). As reported, miRNA isoforms (isomiRs) (Pritchard et al., 2012), which were generated by the addition or deletion of one or multiple nucleotides as terminal modifications, have been shown to participate in proliferative diseases (Boele et al., 2014) and cancer (Wu et al., 2015). On the other hand, N6-methyl-adenosine (m6A) modification, which serves as a RNA epigenetic marker (Niu et al., 2013) analogous to DNA methylation and histone modification (Liu and Pan, 2015), plays important roles in
miRNA biogenesis (Alarcon et al., 2015; Berulava et al., 2015) and is related to the development of cancer (Konno et al., 2019). These functional roles of miRNAs are associated with their specific sequences and chemical modifications. However, it is a challenge to directly probe this information using existing tools for miRNA characterization.

Conventionally, miRNAs can be characterized by northern blot, quantitative reverse transcription real-time polymerase chain reaction (qRT-PCR) assays or microarrays (Pritchard et al., 2012). Other emerging platforms for miRNA sensing include colorimetry (Wu et al., 2016), bioluminescence (Xu et al., 2017), enzymatic activity (Li et al., 2016), nanopore assays (Wang et al., 2011) and electrochemistry (Kilic et al., 2018). Unfortunately, these sensing methods provide limited analytical information because miRNA sequences are not directly reported and prior knowledge of the target miRNA sequence is required.

Next generation sequencing (NGS) (Pritchard et al., 2012), which precisely identifies miRNAs without a prior need of its sequence information, is most widely used in miRNA investigations. However, existing approaches of miRNA sequencing by NGS are carried out by performing reverse transcription followed with deep sequencing of its complementary DNAs (cDNA) (Creighton et al., 2009), which suffers from unpredictable amplification biases and the loss of epigenetic information (Ozsolak and Milos, 2011; Pritchard et al., 2012). At present, there is no sequencing platforms that can directly and simultaneously report miRNA sequence and its base modifications along with a single molecule resolution.

Recent developments in nanopore sequencing have demonstrated direct sequence readout from long stretches of DNA (Laszlo et al., 2014; Manrao et al., 2012) or RNA (Garalde et al., 2018) with a single molecule resolution and could directly resolve base modifications (Wang et al., 2019a; Wescoe et al., 2014). Presumably limited by the existing nanopore sequencing configuration and a lack of compatible motor proteins,
direct sequencing of miRNAs and other short nucleic acid strands has unfortunately never been reported.

Nanopore Induced Phase-Shift Sequencing (NIPSS) (Yan et al., 2019), which is a variant of nanopore sequencing, was recently developed as a universal strategy to sequence analytes other than long stretches of DNA (Laszlo et al., 2014) or RNA (Garalde et al., 2018), such as 2'-deoxy-2'-fluoroarabinonucleic acid (FANA) or other Xeno Nucleic Acids (Yan et al., 2019). However, NIPSS is currently limited by a short read-length of ~15 nucleotides due to the length of the biological nanopore being used (Yan et al., 2019). Mature miRNAs, measuring ~22 nucleotides (nt) in length, are ideal analytes for NIPSS, even with its current configuration. Though the 15 nt read-length fails to cover the miRNA length completely, this technique is superior in principle to existing miRNA sensing methods because no amplification or prior knowledge of target sequences is required, while the epigenetic information within the sequence is still retained.

RESULTS AND DISCUSSION

Direct miRNA sequencing using NIPSS:

To perform direct miRNA sequencing using NIPSS, the target miRNA strand must be conjugated with a section of DNA to form a DNA-miRNA chimeric template. In the eventual configuration, this chimera can be formed via chemical or enzymatic ligation, of which the miRNA section could be a strand of natural miRNA or its synthetic equivalent. However, to quickly prove this feasibility, this chimeric template, which is composed of a segment of DNA on the 3'-end and a segment of miRNA on the 5'-end with the two segments separated by an abasic spacer, was completely custom synthesized for a proof of concept demonstration (Fig. 1a, Table S1). The abasic spacer serves to report a unique signal during NIPSS indicating a clear transition from reading DNA to miRNA. The sequencing library was constructed by thermal annealing from three separate strands: the chimeric template, the primer and the blocker (Fig. 1a, Table S1, Supplemental Information – SI Methods) (Cherf et al., 2012; Yan et al., 2019).
NIPSS was carried out with a mutant *Mycobacterium smegmatis* porin A (MspA) nanopore (SI Methods), atop of which a wildtype (WT) phi29 DNA polymerase (DNAP) served as the ratcheting enzyme during sequencing (Yan et al., 2019). During NIPSS (SI Methods), the sequencing library, which is bound with the phi29 DNAP, was first electrophoretically dragged into the nanopore so that the blocker strand is mechanically unzipped. Subsequently, a phi29 DNAP-driven primer extension was initiated so that the chimeric template starts moving against the electrophoretic force in steps equivalent to a single nucleotide (Fig. S1). The MspA nanopore reports a simultaneous reading of 4 nucleotides around the vicinity of its pore restriction.

As presented in Fig. 1b, the DNA segment, the abasic spacer and the miRNA segment sequentially pass through the nanopore constriction during NIPSS. The DNA segment thus acts as the “DNA drive strand”. Utilizing the phase-shift between the polymerase synthesis site and the nanopore constriction, the phi29 DNAP enzymatically drives the DNA segment to move against the electrophoretic force and simultaneously the tethered miRNA segment is sequenced by the nanopore constriction. The DNA segment is designed to contain sequence repeats between “AAGA” and “TTTC” (3’-5’) to generate a unique signal pattern during NIPSS (Fig. S2), which marks the initiation of a NIPSS event. The abasic spacer “X” following the DNA segment is expected to produce a high current signature that marks the initiation of miRNA sequencing within a NIPSS event (Fig. S1, Fig. S2, Table S1). Since this initiation, miRNA sequencing signals with a read-length of ~15 nucleotides are expected until the abasic spacer reaches the polymerase synthesis site (Yan et al., 2019). Unless otherwise stated, all NIPSS assays described in this paper were carried out by following this configuration. However, the miRNA segments from different chimeric templates contain varying sequences, whereas the DNA segment and the abasic spacer are kept unchanged (Table S1). It should be noted that even without the abasic spacer, we can still in principle discriminate miRNA signals during NIPSS. However, the placement of this abasic spacer serves as a clear signal marker, perfectly suited for a proof of concept demonstration for now.

Experimentally, miR-21, which is an intensively studied oncogenic miRNA and a cancer
biomarker (Chan et al., 2005), was included in the miRNA segment of a chimeric strand (DNA-miR-21, Table S1) for sequencing. A representative raw current trace of DNA-miR-21 is shown in Fig. 1c. This trace can be segmented according to the characteristic sequencing pattern of the DNA drive strand and the abasic spacer respectively, from which the DNA segment reports two repeats of triangular shaped signals and the abasic spacer reports an abnormally high step immediately after the signal from the DNA (Fig. S2). Sequencing signals immediately subsequent from the abasic spacer report sequencing signals for miRNA. Specifically, the representative trace in Fig. 1c reports sequencing signals from miR-21.

Mean current values of signal plateau during NIPSS were extracted from raw sequencing traces to form a squiggle plot (SI Methods, Fig. S3), of which each step represents a simultaneous reading of 4 nucleotides. Here, the squiggle plot (van Dijk et al., 2018), which time-normalizes each signal plateau from raw sequencing traces, is useful in the demonstration of signal consistency between independent events. Squiggle plots from 24 independent sequencing events acquired from DNA-miR-21 were overlapped in Fig. 1d, from which the characteristic high and low current steps which correspond to nanopore readings of AAGA and TTTC were clearly recognized. After the signal from the abasic spacer, 14 subsequent steps that correspond to the sequence of miR-21 were demonstrated. Thus, direct miRNA sequencing by NIPSS has been conceptually demonstrated by sequencing a synthetic miR-21 moiety in the form of a DNA-miRNA chimera. NIPSS events from the same chimeric template demonstrate highly consistent nanopore sequencing patterns, which can be aligned with designed sequences.

The 2-8 nucleotides at the 5’ end of mature miRNAs serve as their seed sequence and normally report similar sequences among miRNAs from the same family. Whereas, miRNAs differ more significantly in sequence at their 3’ end. By analyzing all mature human miRNAs, as downloaded from the miRbase (http://www.mirbase.org/), it has been evidenced that ~99.5% of human miRNAs are in principle distinguishable by sequencing the first 15 nucleotides at their 3’ end (Fig. S4, Table S2). These results indicate that
NIPSS may already serve as a practical method for identifying other miRNAs, even with its current configuration.

**Discrimination of miRNA identities by NIPSS**

Let-7a, which is a member of let-7 family miRNA (Pasquinelli et al., 2000), was included to form another chimeric template named DNA-let-7a for sequencing. Let-7 miRNA was the first human miRNA to have been discovered (Pasquinelli et al., 2000). In contrast to miR-21, which is a cancer biomarker, let-7 miRNA is known to target many oncogenes and thus it behaves as a cancer suppressor (Esquela-Kerscher and Slack, 2006). Simultaneous discrimination of miR-21 and let-7a, which are an oncogenic miRNA and a cancer suppressor miRNA respectively, thus show significant bioanalytical value for cancer diagnosis and serve as an excellent example of miRNA identity recognition by NIPSS (Chen and Qin, 2011; Shishodia et al., 2015).

NIPSS sequencing of DNA-let-7a was carried out in a manner similar to that demonstrated in Fig. 1. Representative raw sequencing data for DNA-miR-21 and DNA-let-7a are demonstrated in Fig. S5. An overlay of squiggle plots from 24 independent DNA-miR-21 sequencing events (Fig. 2a) and 12 independent DNA-let-7a sequencing events (Fig. 2b) are summarized. The means and standard deviations from the extracted events in Fig. 2a and 2b are shown together in Fig. 2c. The sequencing steps from both chimeric templates show great alignment in the DNA part of the events. This is expected because the sequence of the DNA segment is identical in DNA-miR-21 and DNA-let-7a (Table S1). Starting from the steps of the abasic spacer however, the sequencing steps appear to deviate from one another due to sequence variations in their miRNA segments. Though preliminary, the observed variations of sequencing signals between let-7a and miR-21 support the hypothesis that different miRNA identities can be identified from a NIPSS assay.

To demonstrate that, a mixture of DNA-miR-21 and DNA-let-7a was subsequently sequenced using NIPSS. From measurements with the same nanopore, sequencing events of both chimeras were respectively observed, from which signatures of miR-21 or
let-7a were quantitatively confirmed according to their signal variations from the aforementioned squiggle plot results when DNA-miR-21 or DNA-let-7a were respectively measured (Fig. S6). This result thus successfully demonstrates that NIPSS is able to identify different miRNAs from their mixed form, especially when they show a major variation in their sequence, which is demonstrated with miR-21 and let-7a as a proof of concept.

**Discriminating miRNA isoforms by NIPSS:**

According to the scheme of direct miRNA sequencing by NIPSS, minor sequence variations such as addition, deletion or substitution of nucleotides near the 3’-end of the target miRNA could in principle be resolved by resolution of a single nucleotide. Natural miRNA length isoforms (isomiRs), which result mainly from terminal additions or deletions of an indefinite number of nucleotides at the post-transcriptional level, differ at their 3’ or 5’-end in many mature miRNAs (Neilsen et al., 2012). IsomiRs have attracted significant attention due to their functional roles in diverse biological processes, which include modulation of miRNA stabilization (Boele et al., 2014), regulation of mRNA-targeting efficacy (Wu et al., 2015) and their correlations to different disease states (Boele et al., 2014; Kaushik et al., 2015; Pritchard et al., 2012; Wu et al., 2015). Most isomiRs differ by only a single nucleotide at either the 3’ or the 5’-end, such as mono-uridylation (Fei et al., 2018) or mono-adenylation (Thornton et al., 2014), and thus discrimination by conventional miRNA sensing routines is challenging. However, isomiRs differing at their 3’-ends as a result of non-template enzymatic additions (Koppers-Lalic et al., 2014) are more frequently observed and can be immediately distinguished with the current NIPSS scheme.

The 3’ mono-uridylation product of miR-21, which is an isomiR of miR-21, is named miR-21+U in this paper (Fig. 3a). As reported, natural miR-21+U isomiR is abundant in human urine samples and is a promising biomarker for prostate cancer (Koppers-Lalic et al., 2016). As a proof of concept, chimeric template DNA-miR-21 and DNA-miR-21+U (Table S1) were designed and sequenced using the same NIPSS configuration (Fig. 1, 2).
The means and standard deviations of extracted sequencing steps from 24 independent NIPSS events of DNA-miR-21 and DNA-miR-21+U are shown in **Fig. 3b**. As expected, the DNA segment of the NIPSS sequencing signal shows perfect alignment, since the DNA drive strands from both chimeric templates are identical in sequence. Distinct variation of the signal begins at step 23 as a consequence of the additional uridine at the 3’-end of miR-21+U when compared to its isoform miR-21. Starting from step 23, all follow-up sequencing steps from miR-21+U appear back shifted by 1 nucleotide (**Fig. 3b, c**), which further confirms that the 3’ uridylation isomiR has been recognized from direct NIPSS sequencing.

The demonstrated sequencing results between DNA-miR-21 and DNA-miR-21+U have verified that isomiRs with minor sequence variations at the 3’-end can be clearly resolved using NIPSS. As demonstrated, addition or deletion of one or multiple nucleotides can be immediately detected from the characteristic signal variations and the subsequent pattern shift (**Fig. 3c**). More representative data (**Fig. S7**) and detailed statistics (**Fig. S8**) are shown in the Supplemental Information (SI). Minor sequence variations within the first 14 nucleotides to the 3’-end of the target miRNAs can be discriminated by following the same strategy. This makes direct miRNA sequencing by NIPSS immediately applicable to the discrimination between miRNAs from the same family, which share the same 5’ terminal seed sequences but show subtle variations in the sequence region to their 3’ end (Brancati and Großhans, 2018; Moore et al., 2015).

**Detecting m6A modification within miRNA:**

RNA modifications represent a type of epigenetic regulation of gene expression (Lence et al., 2016; Liu and Pan, 2015), of which N6-methyladenosine (m6A) is the most abundant modifications (Lence et al., 2016; Liu and Pan, 2015; Niu et al., 2013). The reversible m6A modification is recognized as a novel epigenetic marker that plays a critical role in mRNA metabolism (Dai et al., 2018), cellular functions (Lence et al., 2016; Roundtree et al., 2017; Zhao et al., 2017) and miRNA biogenesis (Alarcon et al., 2015; Berulava et al., 2015). Recent studies reveal that m6A also exist in mature miRNAs (Berulava et al., 2015;
Konno et al., 2019), and their aberrant level of existence are related to gastrointestinal cancers (Konno et al., 2019). However, due to the chemical and biochemical similarities between adenosine and m6A, precise localization of m6A modifications from natural miRNAs becomes technically challenging for any next generation sequencing platform. The emergence of the technique of MeRIP-seq (methylated RNA immunoprecipitation followed by sequencing) (Dominissini et al., 2012; Meyer et al., 2012) is a major advance for this technical need. However, MeRIP-seq still requires reverse transcription followed by amplification and fails to achieve single nucleotide resolution. Recent developments of direct RNA sequencing using nanopores have successfully demonstrated m6A mapping directly from mRNA samples (Garalde et al., 2018). However, direct RNA sequencing by nanopores is only applicable to sequencing long stretches of RNA instead of miRNAs, whereas MeRIP-seq fails to achieve single molecule and single nucleotide resolutions.

NIPSS is particularly useful in distinguishing minor sequence variations approaching the 3’-end of the target miRNAs. To verify whether single m6A modification within a short stretch of miRNA could be identified by NIPSS, a conceptual experiment was designed. The first nucleotide at the 3’-end of the miRNA segment of DNA-miR-21 is changed from a canonical adenosine to m6A (Fig. 4a). This new strand is named as DNA-miR-21 (m6A) and is custom synthesized for downstream NIPSS characterization. According to the NIPSS convention, the m6A reaches the pore constriction first (Fig. 4b), which makes this nucleotide highly resolvable by the current NIPSS configuration and thus becomes an optimum option for a proof of concept demonstration.

Representative NIPSS sequencing traces from DNA-miR-21 (Fig. 4c) and DNA-miR-21(m6A) (Fig. 4d) show noticeable differences in the current height of the sequencing steps which correspond to the first nucleotide at the 3’-end of the miRNA segment. For a zoomed-in demonstration, only the fraction of the sequencing traces that correspond to the region of interest is shown (Fig. 4c, d) and the additional raw sequencing events are illustrated in Fig. S9. The means and standard deviations of the extracted current steps of 24 independent NIPSS events from DNA-miR-21 (blue) and DNA-miR-21(m6A) (red) are summarized and superimposed in Fig. 4e. Due to the limited
spatial resolution of the MspA constriction, ~3 steps show significant signal variations between the two analytes due to the introduced m6A modification, whereas other sequencing steps show a clear alignment as the remainder of the sequence is identical. The difference of sequencing step heights between DNA-miR-21 (m6A) and DNA-miR-21 are shown in Fig. 4f, from which the maximum signal variation between the two analytes could achieve a difference of ~4 pA in the corresponding positions of the signals, due to the introduced chemical modification (Fig. S10).

**Library preparation by enzymatic conjugation:**

To sequence miRNAs from natural resources by NIPSS, the target miRNA strands must be conjugated chemically or enzymatically with the DNA drive-strand ahead of the library preparation. An enzymatic ligation strategy has been reported (Lee and Yi, 2014) and could be adapted for this purpose.

Briefly, the 5’ phosphorylated DNA drive strand (5PO₄ DNA) is first treated with a 5’ DNA adenylation kit (New England Biolabs), assisted with the Mth RNA ligase (Fig. S11). The adenylated DNA drive strand (5AppDNA) was characterized and purified by ethanol precipitation and further ligated to the 3’-end of the target miRNA strands by the T4 RNA ligase 2 truncated K227Q mutant (T4 Rnl2tr). After this ligation, the DNA-miRNA chimeric template can be characterized by electrophoresis on a 15% polyacrylamide gel and the ligated chimeric template is shown as the extra band of higher molecular weight (Fig. S11).

We then performed NIPSS to sequence these enzymatically ligated chimeric strands. As shown in Fig. S12, sequencing signals from the synthetic and the ligated chimera show strong resemblance in their signal patterns. The means and standard deviations of sequencing steps were extracted from more than 20 independent events from each analyte. These results confirmed that the above demonstrated NIPSS strategy could be applied to sequence miRNAs from natural sources with a current demonstration using their synthetic equivalent.
Conclusion

In summary, we demonstrate single molecule sequencing of miRNA, using a NIPSS approach. Similar strategies could also be adapted to sequence other short non-coding strands, such as siRNA (Bernstein et al., 2001) or piRNA (Siomi et al., 2011), avoiding the laborious motor protein engineering for nanopore sequencing (Garalde et al., 2018). Though demonstrated as a prototype, the NIPSS results in this paper show clear signal discriminations between different sequences, isoforms and epigenetic modifications among synthetic miRNA sequences. MiRNAs from natural resources, such as miRNA extracts from clinical samples, can be conjugated with pre-designed DNA linker strands by performing routine enzymatic ligation to build NIPSS sequencing libraries (Fig. S13). Consequently, direct miRNA sequencing by NIPSS could be directly implemented in clinical diagnosis or could be utilized as a complement to existing miRNA sensing platforms when single base resolution becomes critical. MiRNA sequencing by NIPSS shares the same advantages of other nanopore sequencing technologies, including low cost, single base resolution and portability. In principle, NIPSS, when properly engineered could also be adapted to commercial nanopore sequencers, such as MinION™ (Oxford Nanopore Technologies, UK) (Garalde et al., 2018). However, direct miRNA sequencing has not been demonstrated by Oxford Nanopore Technologies, likely due to the short length of a miRNA strand or its incompatibility with the helicase motor protein. Ultimately, high-throughput direct miRNA sequencing by NIPSS can be carried out in optical nanopore chips (Huang et al., 2015; Wang et al., 2019b) for low cost, high-throughput and multiplexed miRNA characterizations in a disposable device form.

Limitations of the Study

MiRNA sequencing by NIPSS is currently limited to sequence the first 14-15 nucleotides to the 3'-end of any miRNA types. To further extend the read-length of NIPSS to the full length of miRNA, an engineered MspA or phi29 DNAP with redundant structures could be constructed to extend the phase-shift distance of NIPSS. Alternatively, existing nanopores of larger dimensions such as ClyA (Soskine et al., 2012) and FraC (Huang et al., 2017) may be adapted to sequence full length miRNA in its precursor or primary form. Chemical
ligations (Paredes et al., 2011; Vogel and Richert, 2012) between the 5’-end of target miRNAs and the 5’-end of the DNA drive strand may be carried out to form a reverse chimeric strand with a “head to head” configuration for 5’-end miRNA sequencing by NIPSS.

SUPPLEMENTAL INFORMATION

Details of methods, figures and tables are summarized in the separate supplemental information (SI) PDF.

ACKNOWLEDGMENTS

This project was funded by National Natural Science Foundation of China (No. 91753108, No. 21675083, No. 31972917), Fundamental Research Funds for the Central Universities (No. 020514380174, No. 020514380142), State Key Laboratory of Analytical Chemistry for Life Sciences (No. 5431ZZXM1902, No. 5431ZZXM1804), 1000 Plan Youth Talent Program of China, Programs for high-level entrepreneurial and innovative talents introduction of Jiangsu Province. Technology innovation fund program of Nanjing University. Excellent Research Program of Nanjing University (Grant No. ZYJH004).

AUTHOR CONTRIBUTIONS

J.Y.Z. and S.H. conceived the project design. J.Y.Z performed the experiments. S.H.Y. prepared the MspA nanopore. Y.W. designed the DNA sequences. L.C., W.M.G., Y.W., J.Y.Z. and S.H. analyzed the data. Y.Q.W., P.K.Z. and S.H.Y. set up the instruments. S.H. and J.Y.Z. wrote the paper. S.H. and H.Y.C. supervised the project.

DECLARATION OF INTERESTS

S.H., J.Y.Z. and S.H.Y. are inventors on a filed PCT patent application related to this work (PCT/CN2019/102162, Aug/23/2019). The authors declare no other competing interests.

REFERENCES

Alarcon, C.R., Lee, H., Goodarzi, H., Halberg, N., and Tavazoie, S.F. (2015).
N6-methyladenosine marks primary microRNAs for processing. Nature 519, 482-485.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409, 363-366.

Berulava, T., Rahmann, S., Rademacher, K., Klein-Hitpass, L., and Horsthemke, B. (2015). N6-adenosine methylation in MiRNAs. PLoS One 10, e0118438.

Boele, J., Persson, H., Shin, J.W., Ishizu, Y., Newie, I.S., Sokilde, R., Hawkins, S.M., Coarfa, C., Ikeda, K., Takayama, K., et al. (2014). PAPD5-mediated 3' adenylation and subsequent degradation of miR-21 is disrupted in proliferative disease. Proc. Natl. Acad. Sci. USA 111, 11467-11472.

Bracken, C.P., Scott, H.S., and Goodall, G.J. (2016). A network-biology perspective of microRNA function and dysfunction in cancer. Nat. Rev. Genet. 17, 719-732.

Brancati, G., and Großhans, H. (2018). An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity. Nucleic Acids Res. 46, 3259-3269.

Chan, J.A., Krichevsky, A.M., and Kosik, K.S. (2005). MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. Cancer Res. 65, 6029-6033.

Chen, X., and Qin, Z. (2011). Post-transcriptional regulation by microrna-21 and let-7a microRNA in paediatric cholesteatoma. J. Int. Med. Res. 39, 2110-2118.

Cherf, G.M., Lieberman, K.R., Rashid, H., Lam, C.E., Karplus, K., and Akeson, M. (2012). Automated forward and reverse ratcheting of DNA in a nanopore at 5-A precision. Nat. Biotechnol. 30, 344-348.

Creighton, C.J., Reid, J.G., and Gunaratne, P.H. (2009). Expression profiling of microRNAs by deep sequencing. Briefings Bioinf. 10, 490-497.

Dai, D., Wang, H., Zhu, L., Jin, H., and Wang, X. (2018). N6-methyladenosine links RNA metabolism to cancer progression. Cell Death Dis. 9, 124.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485, 201-206.

Dong, H., Lei, J., Ding, L., Wen, Y., Ju, H., and Zhang, X. (2013). MicroRNA: function, detection, and bioanalysis. Chem. Rev. 113, 6207-6233.
Esquela-Kerscher, A., and Slack, F.J. (2006). Oncomirs—microRNAs with a role in cancer. Nat. Rev. Cancer 6, 259-269.

Fei, Q., Yu, Y., Liu, L., Zhang, Y., Baldrich, P., Dai, Q., Chen, X., and Meyers, B.C. (2018). Biogenesis of a 22-nt microRNA in Phaseoleae species by precursor-programmed uridylation. Proc. Natl. Acad. Sci. USA 115, 8037-8042.

Garalde, D.R., Snell, E.A., Jachimowicz, D., Sipos, B., Lloyd, J.H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A., et al. (2018). Highly parallel direct RNA sequencing on an array of nanopores. Nat. Methods 15, 201-206.

Hatziapostolou, M., Polytarchou, C., Aggelidou, E., Drakaki, A., Poultsides, George A., Jaeger, Savina A., Ogata, H., Karin, M., Struhl, K., Hadzopoulou-Cladaras, M., et al. (2011). An HNF4α-miRNA Inflammatory Feedback Circuit Regulates Hepatocellular Oncogenesis. Cell 147, 1233-1247.

He, Y., Lin, J., Kong, D., Huang, M., Xu, C., Kim, T.K., Etheridge, A., Luo, Y., Ding, Y., and Wang, K. (2015). Current State of Circulating MicroRNAs as Cancer Biomarkers. Clin. Chem. 61, 1138-1155.

Huang, G., Willems, K., Soskine, M., Wloka, C., and Maglia, G. (2017). Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores. Nat. Commun. 8, 935.

Huang, S., Romero-Ruiz, M., Castell, O.K., Bayley, H., and Wallace, M.I. (2015). High-throughput optical sensing of nucleic acids in a nanopore array. Nat. Nanotechnol. 10, 986-991.

Kaushik, A., Saraf, S., Mukherjee, S.K., and Gupta, D. (2015). miRMOD: a tool for identification and analysis of 5’ and 3’ miRNA modifications in Next Generation Sequencing small RNA data. PeerJ 3, e1332.

Kilic, T., Erdem, A., Ozsoz, M., and Carrara, S. (2018). microRNA biosensors: Opportunities and challenges among conventional and commercially available techniques. Biosens. Bioelectron. 99, 525-546.

Kim, J., Yao, F., Xiao, Z., Sun, Y., and Ma, L. (2018). MicroRNAs and metastasis: small RNAs play big roles. Cancer Metastasis Rev. 37, 5-15.

Konno, M., Koseki, J., Asai, A., Yamagata, A., Shimamura, T., Motooka, D., Okuzaki, D., Kawamoto, K., Mizushima, T., and Eguchi, H. (2019). Distinct methylation levels of mature microRNAs in gastrointestinal cancers. Nat. Commun. 10, 3888.
Koppers-Lalic, D., Hackenberg, M., Bijnsdorp, I.V., van Eijndhoven, M.A.J., Sadek, P., Sie, D., Zini, N., Middeldorp, J.M., Ylstra, B., de Menezes, R.X., et al. (2014). Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. Cell Rep. 8, 1649-1658.

Koppers-Lalic, D., Hackenberg, M., De Menezes, R., Misovic, B., Wachalska, M., Geldof, A., Zini, N., De Reijke, T., Wurdinger, T., and Vis, A.J.O. (2016). Non-invasive prostate cancer detection by measuring miRNA variants (isomiRs) in urine extracellular vesicles. Oncotarget 7, 22566-22578.

Laszlo, A.H., Derrington, I.M., Ross, B.C., Brinkerhoff, H., Adey, A., Nova, I.C., Craig, J.M., Langford, K.W., Samson, J.M., Daza, R., et al. (2014). Decoding long nanopore sequencing reads of natural DNA. Nat. Biotechnol. 32, 829-833.

Lee, J.E., and Yi, R. (2014). Highly efficient ligation of small RNA molecules for microRNA quantitation by high-throughput sequencing. J. Vis. Exp. 93, e52095.

Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M., et al. (2016). m(6)A modulates neuronal functions and sex determination in Drosophila. Nature 540, 242-247.

Li, B., Liu, F., Peng, Y., Zhou, Y., Fan, W., Yin, H., Ai, S., and Zhang, X. (2016). Two-stage cyclic enzymatic amplification method for ultrasensitive electrochemical assay of microRNA-21 in the blood serum of gastric cancer patients. Biosens. Bioelectron. 79, 307-312.

Liu, N., and Pan, T. (2015). RNA epigenetics. Transl. Res. 165, 28-35.

Manrao, E.A., Derrington, I.M., Laszlo, A.H., Langford, K.W., Hopper, M.K., Gillgren, N., Pavlenok, M., Niederweis, M., and Gundlach, J.H. (2012). Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. Nat. Biotechnol. 30, 349-353.

Mehta, A., and Baltimore, D. (2016). MicroRNAs as regulatory elements in immune system logic. Nat. Rev. Immunol. 16, 279-294.

Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3’ UTRs and near stop codons. Cell 149, 1635-1646.

Moore, M.J., Scheel, T.K., Luna, J.M., Park, C.Y., Fak, J.J., Nishiuchi, E., Rice, C.M., and Darnell, R.B. (2015). miRNA–target chimeras reveal miRNA 3’-end pairing as a major determinant of Argonaute target specificity. Nat. Commun. 6, 8864.
Nassar, F.J., Nasr, R., and Talhouk, R. (2017). MicroRNAs as biomarkers for early breast cancer diagnosis, prognosis and therapy prediction. Pharmacol. Ther. 172, 34-49.

Neilsen, C.T., Goodall, G.J., and Bracken, C.P. (2012). IsomiRs—the overlooked repertoire in the dynamic microRNAome. Trends Genet. 28, 544-549.

Niu, Y., Zhao, X., Wu, Y.S., Li, M.M., Wang, X.J., and Yang, Y.G. (2013). N6-methyl-adenosine (m6A) in RNA: an old modification with a novel epigenetic function. Genomics, Proteomics Bioinf. 11, 8-17.

Ozsolak, F., and Milos, P.M. (2011). RNA sequencing: advances, challenges and opportunities. Nat. Rev. Genet. 12, 87-98.

Paredes, E., Evans, M., and Das, S.R. (2011). RNA labeling, conjugation and ligation. Methods 54, 251-259.

Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Müller, P., et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408, 86-89.

Pauley, K.M., Cha, S., and Chan, E.K. (2009). MicroRNA in autoimmunity and autoimmune diseases. J. Autoimmun. 32, 189-194.

Pritchard, C.C., Cheng, H.H., and Tewari, M. (2012). MicroRNA profiling: approaches and considerations. Nat. Rev. Genet. 13, 358-369.

Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. Cell 169, 1187-1200.

Shishodia, G., Shukla, S., Srivastava, Y., Masaldan, S., Mehta, S., Bambhani, S., Sharma, S., Mehrotra, R., Das, B.C., and Bharti, A.C. (2015). Alterations in microRNAs miR-21 and let-7a correlate with aberrant STAT3 signaling and downstream effects during cervical carcinogenesis. Mol. Cancer 14, 116.

Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. 12, 246-258.

Soskine, M., Biesemans, A., Moeyaert, B., Cheley, S., Bayley, H., and Maglia, G. (2012). An engineered ClyA nanopore detects folded target proteins by selective external association and pore entry. Nano Lett. 12, 4895-4900.

Teng, Y., Ren, Y., Hu, X., Mu, J., Samykutty, A., Zhuang, X., Deng, Z., Kumar, A., Zhang, L., Merchant, M.L., et al. (2017). MVP-mediated exosomal sorting of miR-193a promotes colon
cancer progression. Nat. Commun. 8, 14448.

Thornton, J.E., Du, P., Jing, L., Sjekloca, L., Lin, S., Grossi, E., Sliz, P., Zon, L.I., and Gregory, R.I. (2014). Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). Nucleic Acids Res. 42, 11777-11791.

van Dijk, E.L., Jaszczyszyn, Y., Naquin, D., and Thermes, C. (2018). The third revolution in sequencing technology. Trends Genet. 34, 666-681.

Vogel, H., and Richert, C. (2012). Labeling small RNAs through chemical ligation at the 5' terminus: enzyme-free or combined with enzymatic 3'-labeling. Chembiochem 13, 1474-1482.

Wang, Y., Patil, K.M., Yan, S., Zhang, P., Guo, W., Wang, Y., Chen, H.Y., Gillingham, D., and Huang, S. (2019a). Nanopore sequencing accurately identifies the mutagenic DNA lesion O6-carboxymethyl guanine and reveals its behavior in replication. Angew. Chem. Int. Ed. 58, 8432-8436.

Wang, Y., Wang, Y., Du, X., Yan, S., Zhang, P., Chen, H.-Y., and Huang, S. (2019b). Electrode-free nanopore sensing by DiffusiOptoPhysiology. Sci. Adv. 5, eaar3309.

Wang, Y., Zheng, D., Tan, Q., Wang, M.X., and Gu, L.Q. (2011). Nanopore-based detection of circulating microRNAs in lung cancer patients. Nat. Nanotechnol. 6, 668-674.

Wescoe, Z.L., Schreiber, J., and Akeson, M. (2014). Nanopores discriminate among five C5-cytosine variants in DNA. J. Am. Chem. Soc. 136, 16582-16587.

Wu, H., Liu, Y., Wang, H., Wu, J., Zhu, F., and Zou, P. (2016). Label-free and enzyme-free colorimetric detection of microRNA by catalyzed hairpin assembly coupled with hybridization chain reaction. Biosens. Bioelectron. 81, 303-308.

Wu, X., Zeng, R., Wu, S., Zhong, J., Yang, L., and Xu, J. (2015). Comprehensive expression analysis of miRNA in breast cancer at the miRNA and isomiR levels. Gene 557, 195-200.

Xu, Q., Ma, F., Huang, S.Q., Tang, B., and Zhang, C.Y. (2017). Nucleic Acid Amplification-Free Bioluminescent Detection of MicroRNAs with High Sensitivity and Accuracy Based on Controlled Target Degradation. Anal. Chem. 89, 7077-7083.

Yan, S., Li, X., Zhang, P., Wang, Y., Chen, H.Y., Huang, S., and Yu, H. (2019). Direct sequencing of 2'-deoxy-2'-fluoroarabinonucleic acid (FANA) using nanopore-induced phase-shift sequencing (NIPSS). Chem. Sci. 10, 3110-3117.
Zhao, B.S., Roundtree, I.A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. Nat. Rev. Mol. Cell Biol. 18, 31-42.
Figure legends

Figure 1. Direct miRNA sequencing using NIPSS.

(a) A schematic diagram of the preparation of a sequencing library. The miRNA sequencing library is thermally annealed (SI Methods) from three separate nucleic acid strands which include a chimeric template, a primer (green) and a blocker (light blue) (Table S1). The chimeric template is composed of a miRNA segment (red), an abasic residue (blue dot) and a DNA segment (black). (b) The NIPSS strategy for direct miRNA sequencing. NIPSS is carried out with an MspA nanopore (purple) and a wildtype (WT) phi29 DNAP (green) by following the reported enzymatic ratcheting strategy (Fig. S1) (Manrao et al., 2012). A fixed phase shift distance between the polymerase synthesis site and the pore constriction is utilized to directly sequence miRNA. During NIPSS, the DNA segment, the abasic residue and the miRNA segment sequentially move through the nanopore constriction in single nucleotide steps until the abasic site (blue dot) reaches the binding pocket of phi29 DNAP. The inset image shows a zoomed-in view of the pore constriction. Due to the limited spatial resolution of the pore constriction, nanopore sequencing signals from MspA results from simultaneous reading of different combinations of sequence quadromers spanning the pore constriction during NIPSS. (c) A typical current trace acquired by sequencing DNA-miR-21 (Table S1) using NIPSS. The DNA and miRNA segments of the trace are marked with black and red lines, respectively. The trace segment that corresponds to reading the abasic site (X) is marked with a blue stripe. Briefly, the DNA segment of the trace appears as two triangular shaped current characteristics due to the sequence design (Fig. S2). The DNA segment of the trace is immediately followed by an abrupt increase of the signal due to the introduction of the abasic site after the DNA sequence. All step transitions after the abasic signal lead to miRNA sequencing signals. (d) Overlay of multiple time-normalized events (N=24) from the DNA-miR-21 results acquired by NIPSS. Each step in a sequencing event was extracted by a level detection algorithm (Methods 4, Fig. S3). The corresponding sequence (3'-5' convention, if not otherwise stated) is aligned above the statistics. The current levels of “AAGA” and “TTTC”, which represents the highest and the lowest sequencing signals by reading the DNA part, are marked by black arrows. The demonstrated results were acquired by performing NIPSS with an aqueous buffer of 0.3 M KCl, 10 mM HEPES/KOH, 10 mM MgCl₂, 10 mM (NH₄)₂SO₄ and 4 mM DTT at pH 7.5. See also Transparent Methods, Table S1 and Figures S1-S3.

Figure 2. Identification of miRNAs using NIPSS.

Chimeric template strands containing miRNA-21 and Let-7a sequences were custom synthesized (Table S1) and directly sequenced by NIPSS. (a) Overlay of multiple time-normalized events (N=24) from the DNA-miR-21 results acquired by NIPSS. (b) Overlay of multiple time-normalized events (N=12) from the DNA-let-7a results acquired by NIPSS. (a-b) The corresponding sequences (DNA-miR-21 or DNA-Let-7a, 3'-5' convention) are aligned above the plots. The DNA segments are marked in black and the miRNA segments are marked in red. The abasic site (X, blue), which separates the DNA and miRNA segments, acts as a signal marker to identify the sequence transition from reading DNA to miRNA during NIPSS. (c) Consensus sequencing results comparison between DNA-miR-21 and DNA-let-7a. The mean and standard deviation values are derived from time-normalized events, as demonstrated in a and b. The DNA part of the NIPSS results shows great alignment in all steps between both templates. However, the miRNA segment of the signals shows significant variations, starting from the step marked with the blue stripe. Blue stripes in (a-c) mark the sequencing step of TCAX, which is the first quadromer sequence containing the abasic residue when acquired by NIPSS. See also Table S1, Figure S5 and Figure S6.
Figure 3. Discrimination of miRNA isoforms (isomiRs) using NIPSS.

(a) miR-21 and its isoforms. The upper image demonstrates the structure of a precursor microRNA (pre-miRNA) for human miR-21. The lower image demonstrates the sequence (5'-3') of mature miR-21 and its miR-21+U isoform, respectively. An additional uracil (red box) exists at the 3'-end of miR-21+U. (b) Comparison of consensus sequencing results between DNA-miR-21 and DNA-miR-21+U. Consensus sequencing results from NIPSS reading of DNA-miR-21 (brown) and DNA-miR-21+U (blue) derived from 24 time-normalized independent events (Table S3). The DNA, the abasic site and the miRNA part of the signal are marked separately. The blue strip marks the sequencing step of TCAX, which is the first sequence quadromer containing an abasic spacer encountered by NIPSS. The demonstrated statistical results show great alignment in all parts except that marked with a dashed-line box. (c) Sequencing result shift between DNA-miR-21 and DNA-miR-21+U. A zoomed-in view of the dashed-line box in b illustrates a shift effect of current levels caused by a nucleotide insertion of DNA-miR-21+U in reference to DNA-miR-21. The aligned sequence context above the results demonstrates that the addition of a uracil (red box) in DNA-miR-21+U systematically generates 1 nucleotide (marked as δ) shift. This single nucleotide result shift is also demonstrated by the schematic diagram in the image inset which takes the results of step 25 as an example. See also Table S1 and Table S3, Figure S7 and Figure S8.

Figure 4. Direct N6-methyladenosine (m6A) mapping using NIPSS.

(a) Schematic diagram of chimeric template strands containing canonical adenosine (A) or N6-methyladenosine (m6A). A single “A” or “m6A” is embedded in different chimeric strands (top: DNA-miR-21, bottom: DNA-miR-21(m6A), Table S1) for NIPSS sequencing, where its chemical structure and location is annotated. Black and red segments represent the DNA and the miRNA part of the strand. (b) Schematic diagram of NIPSS sequencing of miRNA containing “A” or “m6A” nucleotides. A* within the sequence context represents the “A” or “m6A” nucleotide within the strand. (c) A representative current trace of DNA-miR-21 sequenced by NIPSS. Blue lines represent extracted mean current values from each step. The corresponding sequence context is aligned below in which a canonical adenosine is marked in blue. (d) A representative current trace when DNA-miR-21(m6A) is sequenced by NIPSS. Red lines represent extracted mean current values from each step. The corresponding sequence context is aligned below, from which a N6-methyladenosine is marked in red. Scale bars in c and d represents 10 pA (current, vertical) and 50 ms (time, horizontal), respectively. (e) Consensus sequencing results comparison between DNA-miR-21 and DNA-miR-21(m6A). The means and standard deviations were derived from 24 independent events. A* within the sequence context below the results represents either A or m6A. (f) Current differences between NIPSS results of DNA-miR-21 and DNA-miR-21(m6A). These differences were derived by calculating \( \Delta I = I_{DNA-miR-21(m6A)} - I_{DNA-miR-21} \) from the mean values of 24 events from each strand, with the associated sequence aligned below. The standard deviation values of DNA-miR-21 are demonstrated with gray column and the standard deviation of DNA-miR-21(m6A) is demonstrated with black error bars. The m6A modification results in a signal variation when m6A containing sequence quadromers were read by the nanopore constriction. Due to the limited spatial resolution of MspA, a single m6A modification results in detectable signal fluctuations within three current steps as marked by red lines. See also Table S1, Figure S9 and Figure S10.
a
TTCAAGATTCAAGATTTCAAXGUUGUAGUCAGAC

b
TTCAAGATTCAAGATTTCAAXUGAUAGUUGGAU

c
DNA-miR-21
DNA-let-7a
miR-21: UAGCUUAUCAGACUGAUGUUGA

miR-21+U:

a

b

δ

60 
40 
20 
0 10 20 30 
30 
40 
50 
23 25 27 29 31

Step number

20 I (pA) I (pA)

c

GUCGG AGCUUAUC GACUG A UGUUG CUGU G A A

CAGUC UCGGGUAG CUGAC ACAAC G UA C C

U G

U A A A

U

A UG C

A A

U

CU

5’ 3’

CCGUA

DNA miRNA

5 15 25 
33 
U

U

G

U 
A

G 
G

U

U

G 
U

A
Adenosine (A)
N6-methyladenosine (m6A)

DNA-miR-21 (m6A)
Highlights:

- The first demonstration of single molecule miRNA sequencing.
- MiRNA sequencing by NIPSS can directly identify epigenetic modifications.
- Enzymatic conjugation enables NIPSS sequencing of natural miRNAs.