

A DNA circuit that records molecular events

Mingzhi Zhang^{1,3}, Colin Yancey⁴, Chao Zhang¹, Junyan Wang^{1,2}, Qian Ma³, Linlin Yang¹, Rebecca Schulman^{4,5*}, Da Han^{1,2*}, Weihong Tan^{1,2*}

¹Institute of Molecular Medicine (IMM), Shanghai Key Laboratory for Nucleic Acid Chemistry and Nanomedicine, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200127, China.

²Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang 310022, China.

³Intellinosis Biotech Co., Ltd., Shanghai, 201112, China.

⁴Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD 21218, USA.

⁵Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA.

*e-mail: tan@hnu.edu.cn (W.T.), dahan@sjtu.edu.cn (D.H.), rschulm3@jhu.edu (R.S.)

Abstract

In biological systems, transient molecular signals direct the operation of signal transduction and genetic regulatory networks. Characterizing the relative onset times, strength (or concentration) and duration of these signals is critical for understanding the operation of these networks. However, detecting multiple such molecules as they are produced and then quickly consumed is challenging. A molecular event recorder (MER) encodes information about transient molecular events, such as the relative appearance times, concentration, and duration of molecular signals as a stable DNA sequence output amenable to downstream sequencing or other analysis. Here, we report the development of a *de novo* molecular event recorder that processes information using a strand displacement reaction network and produces outputs using the primer exchange reaction (PER). The recorder encodes information about the concentration, duration, of target molecules and the relative order that these molecules appear as the concentrations of different DNA outputs.

Introduction

Molecules regulate biological activities in living cells ¹. The presence of a molecule can represent a temporal molecular event (TME), which can influence or directly indicate a downstream activity's occurrence. One instance of this would be the presence of mRNA in a cell, as it is a TME which leads to the production of a given protein through gene expression, demonstrating the significant effect of TMEs on cell status. The order, magnitude, and duration of the expression of different genes has a profound impact on the evolution of a cell's properties and function ^{2,3}. Some examples include concentrations of mRNA that relate to cytokine expression influencing a cell's differentiation⁴, the length of signal exposures being correlated with the outcome of a natural killer cell's "education" ⁵, which can result in different cell fate ^{6,7}, and in turn affecting homeostasis ⁸. Therefore, an analysis of cells exposed to nucleic acid can yield information about the relative order, initial concentration, and duration of a molecular event, leading to the identification of cell status and development. However, initial exposure information is momentary, and methods aimed at gaining such information in a way that would modify cellular behavior pose risks, the ability to sense multiple properties simultaneously could be lost or

unexpected side effects could occur by the change in target system. Though Molecular event recorders, whose devices can capture and store details for long periods and provide a safe and reliable option for long-term analysis using DNA, are a plausible solution for the capture of molecular data presented momentarily, skilled technicians employing complicated procedures are needed to operate potential molecular event recorders and implement data sensing and recording with engineered enzymes^{9-13 14-16 17 18-21}. To address these challenges, we herein report the design of a molecular event recorder (MER) that processes information using a strand displacement reaction network and produces outputs using the primer exchange reaction (PER)^{22,23}. This type of recorder contains transducers and writers which are used for molecular signals transformation and information recording.

Framework of the MER

Molecular events usually contain information about different types of molecular signals, including the relative order, concentration, and duration. More specifically, relative order refers to the different onset times of target molecules, concentration refers to the initial intensity of the target molecule in observed region and duration refers to the present time of target molecule in this region. Thus, in our framework, Inputs are DNA oligo with specific sequence, and a molecular event can, therefore, be recorded in different domains of a DNA oligo barcode. (Figure 1A). The recorder can be divided into three modules: one that senses and records order; one that senses and records concentration; and one that senses and records duration. A substrate module of time delay purposes is used for long-duration detecting in duration module.

As shown in Figure 1B, reactions among modules are linked, allowing transducers in the concentration and duration modules to respond to the Input and transmit specific signals to activate writer hairpins. Once modules are activated, primers containing a specific order signature start their work (Figure S4). A specific order signature might, for example, be a three-base difference in the primer. By its ability to elongate in writer hairpin using primer exchange reaction (Figure 1C), specific barcode can be continually added from its 3-prime as well as downstream recording primer. As a result, a specific recording product in which information about relative order, type-concentration, and duration is generated.

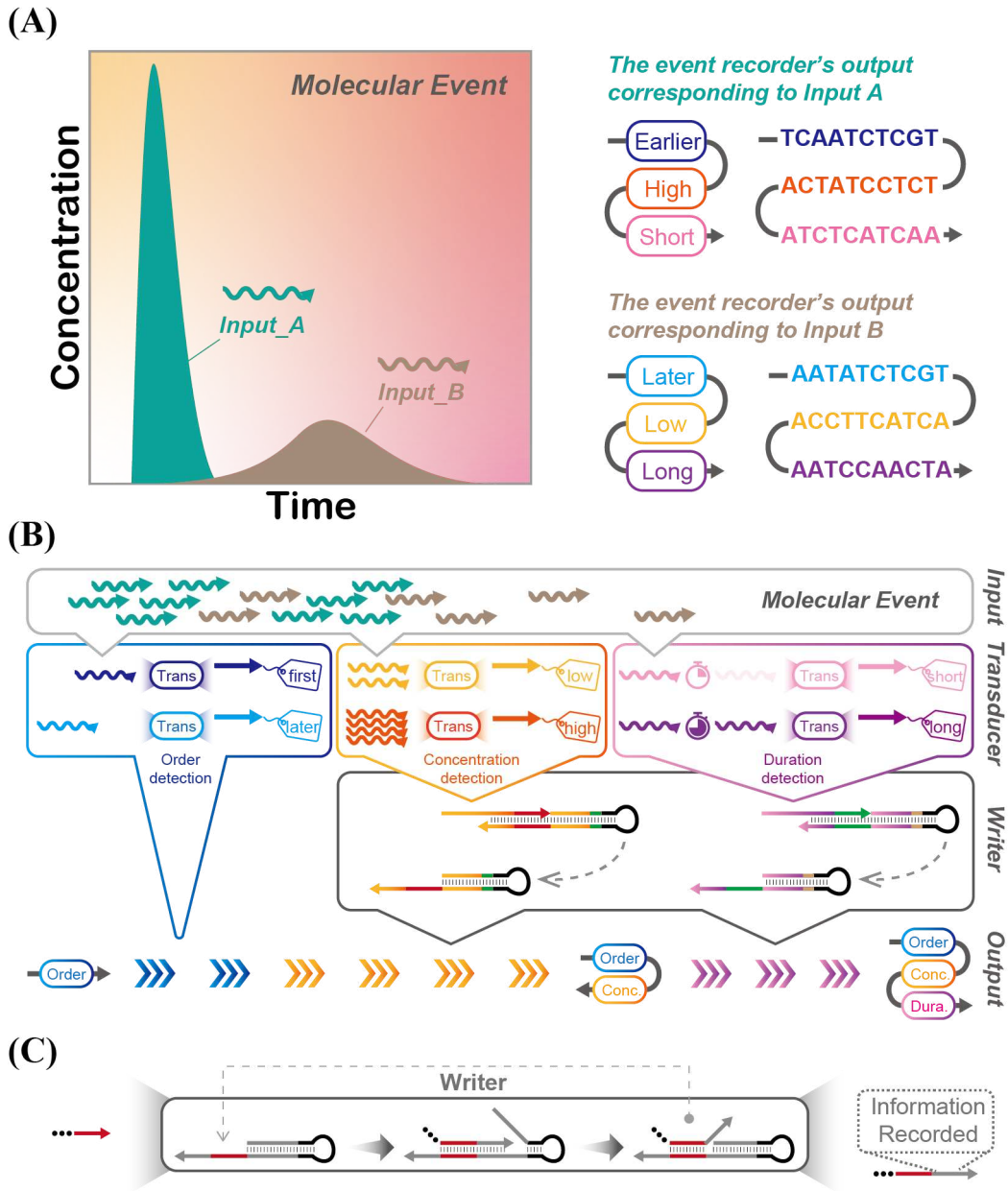


Figure 1. Overview of the molecular event recorder (MER). (A) An example of a molecular event: Input_A appeared with a high concentration and lasted for a brief time, while it was followed by Input_B which appeared with a low concentration and lasted for a long time. (B) Workflow of the recorder. Once the input appears, the transducers of concentration and duration respond by releasing a specific DNA oligo and activating their relevant write hairpins. The recording of information starts from the primer with order identification, and the concentration barcode is elongated by the writer hairpin in the concentration modules and followed by the duration module. (C) Mechanism of primer exchange reaction (PER) used in the writer hairpin.

Relative Order identification

Relative order can be defined as the timing of the molecule's first appearance orderly early or late by time. Usually, more than one molecule takes part in a life activity, a

molecular event always consists of multiple inputs while the relative onset times often lead to distinct consequences. It has been proved that different lengths of the binding domain in primer could cause different reaction priorities on PER hairpin (Figure S27A and S27B) ²³. Thus, we designed two primers with different domain binding lengths (10 nt for p1 and 7 nt for p2) in the concentration module, assuming that p1 would have a higher reaction priority (Figure S4). We used $t(\text{Input}_A) - t(\text{Input}_B)$ to quantify the order and gap time, the first Input appearance might use more p1 to record its information (Figure 2A) and the expectation of [p1_A] and [p1_B] in different order conditions could be that shown like an “X” shape in Figure S6A where the order of Input_A and Input_B significantly result in an increasing ratio of [p1_A] over [p1_B] when $t(\text{Input}_A) - t(\text{Input}_B)$ changed from negative to positive, as shown in Figure S6B.

We modified two different fluorescence on 5' of the primers (6-FAM on p1 and CY5 on p2) to help characterize the consumption of these two primers in one recording system (Figure S7), then tested a list of gap times between Input_A and Input_B through PAGE (Figure 2B). We found Input_A that appeared earlier would cause higher [p1_A] compared to the later appearance condition, while Input_B experienced the same situation (Figure 2C). We plotted [p1_A]/[p1_B] through simulation modeling and determined the threshold value for order identification to be around 2.0 (Figure S28C). We next used a classification model based on this threshold to classify the order of Input_A and Input_B (Figure S6C), and it showed that the monotone change of [p1_A]/[p1_B] helps in correct classification (Figure 2D).

Compared to the goal pattern, [p1_B] in Figure 2C showed a lower yield than expected. We considered a leakage in the inactivated writer hairpin and found it shows higher in A system compared to B system, resulting in a curve shift in the experiment (Supplementary Section 3.1, Figure S9A and S9B).

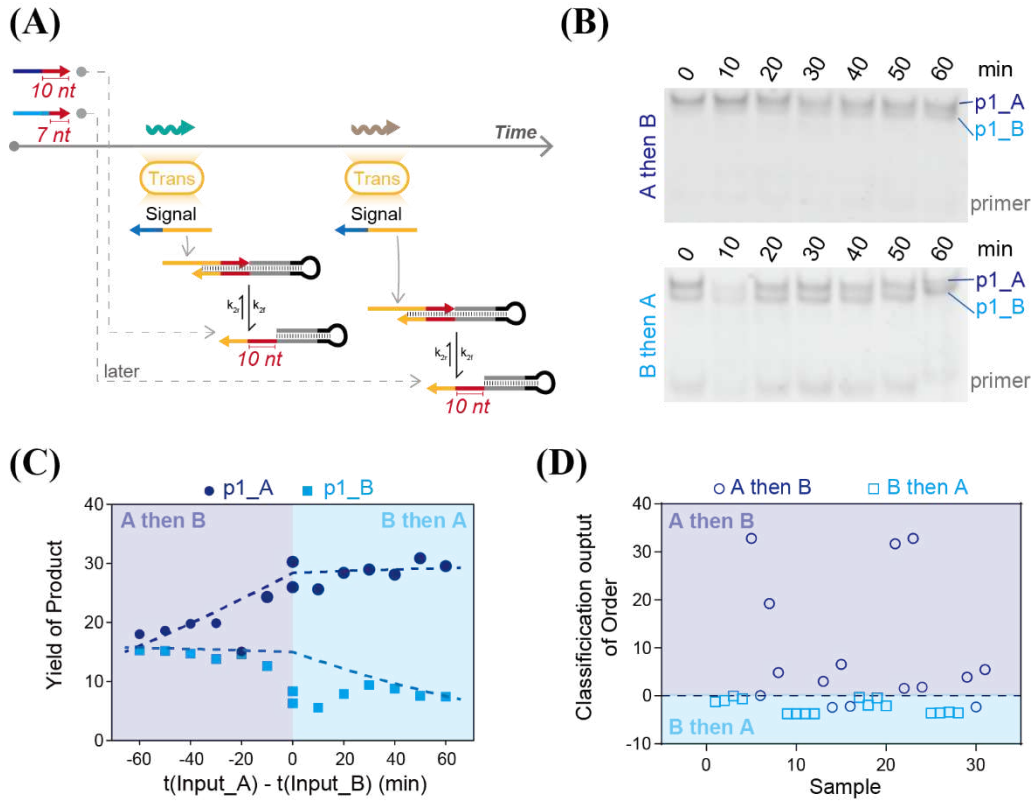


Figure 2. Order identification for double input. (A) Mechanism of order identification. Input activates its specific writer hairpin and consumes the primer with a p1 sign before p2 sign, resulting in using more p1 in the product of the former input. (B) Different input time results in a different yield of products of Input_A (with p1 signature) and products of Input_B (with p1 signature). (C) Yield of the products is plotted under order and interval time of the Input. The solid dots represent the data from PAGE in (B). (D) Scatter distribution of the order. The order of appearance can be divided into two dissimilar parts. Note that $Classification\ output = Ratio - Threshold$.

The Concentration Module in MER

In general, a distinction must be shown between low and high concentration conditions to design a reliable concentration sensing module. A thresholding circuit is able to successfully output a concentration higher than the threshold^{24, 25, 26}. However, to confirm a dependable “low concentration” condition and distinguish it from a “no Input” condition, leakage in strand displacement reactions must be considered. To reduce the influence of leakage, we considered the outputs from the track of both Low and High instead of High only, which works because the influence of leakage would be offset by the comparison of Low and High.

We aimed to achieve the design goal drawn in Figure S10A, which shows a huge yield gap between product_Low and product_High in low-input conditions and a small yield difference in high-input conditions. This difference will lead $[\text{product_Low}]/[\text{product_High}]$ jumping sharply near the critical concentration. In accordance with the expected product yield curves shown in Figure S10B, $[\text{product_Low}]/[\text{product_High}]$ is used to classify low or high concentration of the Input. Transducer_Low, which has a 10 nt toehold (estimated as $3 \times 10^{-3} \text{ nM}^{-1}\text{s}^{-1}$), should therefore react with, as well as consume, the input much faster than Transducer_High which has a 5 nt toehold (estimated as $5 \times 10^{-7} \text{ nM}^{-1}\text{s}^{-1}$)²⁷ (Figure 3A, Supplemental Section 3.2). The resulting curves for the yields of each relative product of these two reactions, when they are in competition, are shown in Figure 3B where yields of the product are analyzed via grey curves. Only product_Low is produced if the initial concentration of the Input is lower than 100 nM, whereas both product_Low and product_High will be produced if the initial concentration of Input is above 200 nM (reaction rate = 10^{-4} s^{-1}).

To understand how the concentration model works, we developed a simulation to determine if increasing or decreasing the rate constants in the concentration module would increase or decrease the final concentration of the products by more than a factor of two, which would indicate a stable prediction model. We found that the rate constants of binding (*primer_hairpin_f*) and releasing (*primer_hairpin_r*) reactions between primer and writer hairpin affect the rate of production and product yield. Moreover, an increasing rate constant of elongation (*primer_hairpin_elongation*) in the writer hairpin would lead to a huge yield gap in high-input conditions. To explain, strand displacement reactions in transducers should be the rate-limiting step in order to show the output difference between low-input and high-input. An inefficient elongation would result in a heavy accumulation of signals from transducers, which unable to distinguish these two situations. However, efficient elongation could lessen the accumulation and reduce the influence at writing step. Consequently, the rate constants discussed above (see Supplementary Section 3.2.2) are significant because they make great effect to the yield of products, which deeply affect the analysis of readouts. These simulated values for *primer_hairpin_r* and *primer_hairpin_elongation* are close to the estimated values of strand releasing and strand elongation reactions, except the value for *primer_hairpin_f*, the binding rate of

which is 30 times slower than the theoretical one, possibly caused by an unexpected secondary structure in the binding site.

According to the simulation results in Figure 3C, low-input conditions (initial concentration lower than 100 nM) showed a large yield difference between product_Low and product_High, while in the conditions of high-input (initial concentration higher than 200 nM), yield of the product was shrinking, thus meeting our expectations. By adjusting the simulation model, decreasing [Transducer_Low] leads to a lower concentration threshold, resulting in the ratio of [product_Low] over [product_High] changed into a slight gradient (Figure S10B and Supplementary Section 3.2).

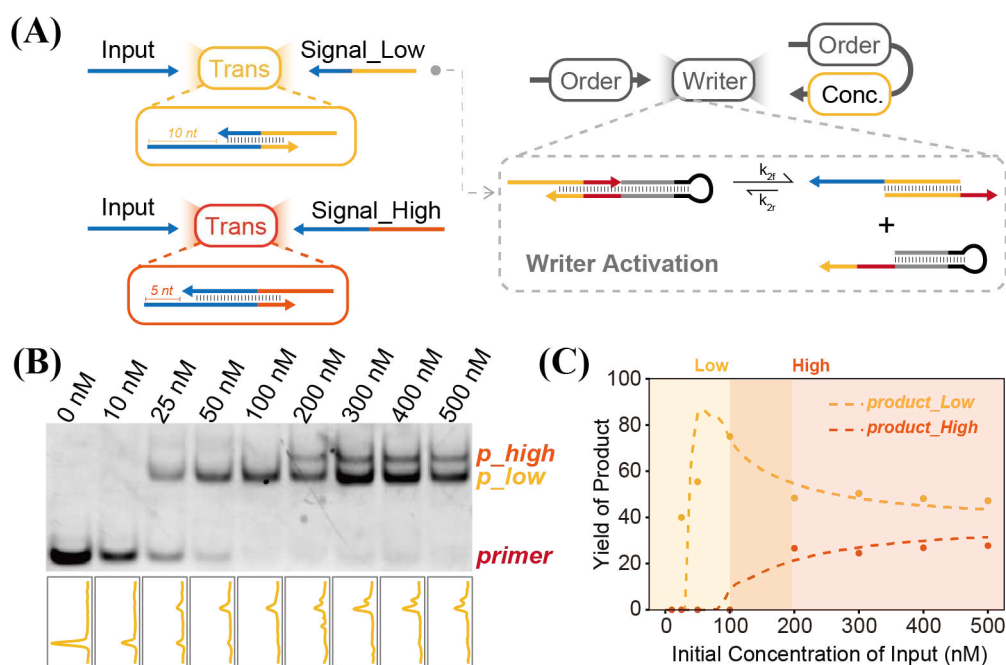


Figure 3. The concentration sensing and recording module. (A) Different lengths of toehold cause distinct reaction rates in toehold-mediated strand displacement reactions. (B) The different initial concentrations of Input result in a different yield of product_Low and product_High. Grey curves show shift of the peak. (C) Yield of the product plotted under initial Input concentration. Solid dots represent the data from PAGE in (B), while the dashed lines come from the simulation.

The Duration Module in MER

The duration module records how long a target molecule is present. We aimed to achieve a huge yield gap between product_Short and product_Long in short duration conditions and an approximate yield in long duration conditions (Figure S16A and

S16B). To help make this decision, we designed a Timer circuit, which is a submodule that produces an output at a fixed time after Input is received. This submodule changes how the circuit operates before and after the delay has elapsed. Several temporal circuits have been designed to delay time, the central idea of which is to “kill time” before releasing the signal. Some of these circuits accumulate intermediate or waiting time until reaching the detecting threshold²⁸, whereas others delay time through barrier consumption, which means signal can be released until the barriers with a higher reaction priority are consumed²⁹⁻³².

In our design, time delay was accomplished through barrier consumption based on PER reaction²⁹, which is easily accommodated by our recording system (Figure 4A). Thus, as soon as the Delay Gate is activated by the Input, it would consume a single-stranded Timer immediately. Since no Delay_Signal is complementary to the Timer, nothing would be released to trigger downstream reaction, so that the Delay Gate achieves time delay. Once the Timer is exhausted, the Timer in the Timer-Delay_Signal complex would start to react, Delay_Signal could be released owing to intramolecular strand displacement in Gate hairpin.

This mechanism makes it possible to adjust the length of time delay such that a higher concentration of Timer would be predicted to produce a longer delay (Figure S16C and Figure 4B). We tested 0×~6× Timer with 1× Timer-Delay_Signal to investigate performance of the delay gate. For each increase of [Timer] by 1× (i.e., 50 nM), we found that the threshold delay time of Delay_Signal increased by nearly one hour. We also designed an AND Gate for Delay_Signal and Input in the duration module and expected that the output Signal_Long only appears after the delay has elapsed (Figure 4A). Figure 4C shows only Input or Delay_Signal present cannot trigger the AND Gate to release Signal_Long until they present together, which matches our expectation.

Transducer_Short should interact with Input as soon as possible to produce product_Short. If the Input is present for only a brief time, product_Short should be the only output of the circuit. If the Input is still present when Delay Signal is released at the Timer’s conclusion, product_Long is also produced, and its proportion of the output increases with increasing Input duration. To test this principle, we measured the response of the circuit to different Input duration times and found that the critical

time of duration at which product_Long starts to be produced is about 60 min (Figure 4D and 4E). Therefore, in the conditions of short duration (present shorter than 30 min), these results show that the reactions achieved a huge yield gap between product_Short and product_Long, whereas they showed an approximate yield in long duration conditions, as expected.

To better understand what controls the dynamics of the duration model, we developed a kinetic reaction model and identified several rate constants that sensitively determine the yield of the recording products. The forward and reverse (or on- and off-) rate constants (k_f and k_r) between the primer and writer hairpin and the effective rate constant for the composite irreversible reaction involving the elongation and strand-displacement release of the recording products are the important parameters. Since the Delay Gate, as well as AND gates, are added as submodules in the duration module, we also found the rate constants of binding between Timer and Gate hairpin ($timer_gate_f$), Timer elongation in Gate hairpin ($timer_gate_elongation$), unbinding between Timer and Delay_Signal ($timer_delay_release$), and activation of the AND Gate by the Delay_Signal ($delayS_to_preSW$) could change the dynamics significantly (Supplementary Section 3.3, Figure S26).

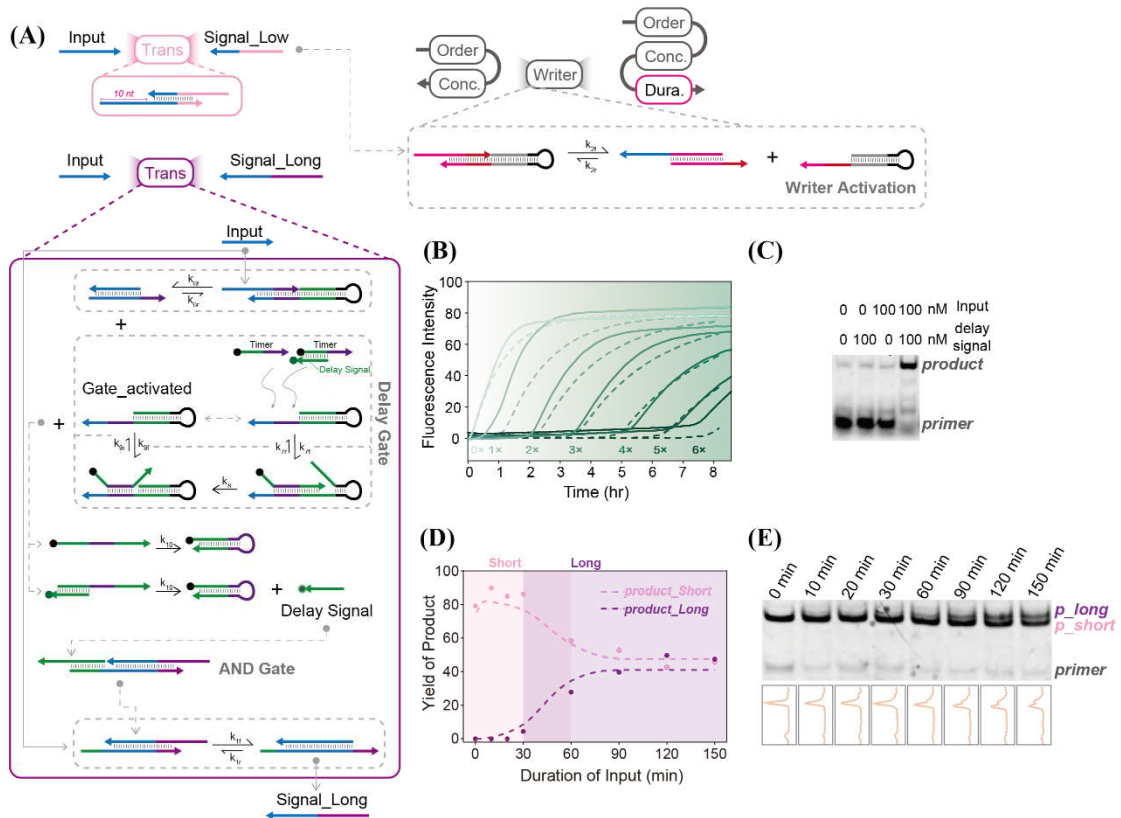


Figure 4. Duration sensing and recording module. (A) Mechanism of duration module. Long duration transducer contains a Delay Gate and an AND Gate. Timer elongation is followed by the activated gate with Input. After the Timer is consumed, it is replaced by Timer_Delay, and the delay signal can be released owing to the intramolecular strand displacement in Gate hairpin and participate in triggering AND Gate for identification of long-time duration. (B) Time-relative fluorescence monitoring of the Timer circuit's output and Delay Signal. Concentration of the Timer-Delay_Signal complex is 50 nM, and concentration of the free Timer varies between 0 to 6 times the concentration of Timer_Delay (50 nM) according to the labels 0 $\times$ -6 $\times$. Dashed lines are from the simulation, and solid lines are from the experiment. (C) PAGE gel showing AND gate function. Only the Delay Signal and the Input appear together, Signal_Long can be released. (D) Yield of the products is plotted under duration time of the Input. Solid dots represent the data from PAGE in (E), while dashed lines come from the simulation. (E) Different duration times of Input result in different yields of product_Short and product_Long. Grey curves are shown to recognize shift of the peak.

Readout and model analysis of molecular events

To record a single input event, concentration and duration modules should be combined. To accomplish this, we designed a primer domain to record duration information in the concentration information writer hairpin (Supplementary Section 1.1.1, Figure S1). We suggested to group and combine the recording products and calculated by assuming that products tuned by the modules separately (Figure 5A). Moreover, we can easily determine the thresholds with monotonicity (Figure 5B and 5C). Plotting the test samples with the classification output of the initial concentration as the abscissa and classification output of duration as the ordinate can easily classify the samples into four different zones (Figure 5D) since the recording products can be both divided into two different parts based on concentration or duration classification. Input_A and Input_B show a similar classification pattern – all the test samples can be classified correctly under concentration, while for duration, “Low-Long” products occupied “Low-Short” zone, which is incorrect. This means that the Input did not successfully activate Transducer_Long, possibly because the Input interacted with surplus Transducer_Low or even Transducer_High faster than with the AND Gate. It is also possible that the concept between “Low” and “Long” might be recognized as a natural contradiction in the recording system. All molecular event test samples are listed in Supplementary Section 2.6, Table S1, and the classification results are shown in Figure 5E for Input_A and 5F for Input_B.

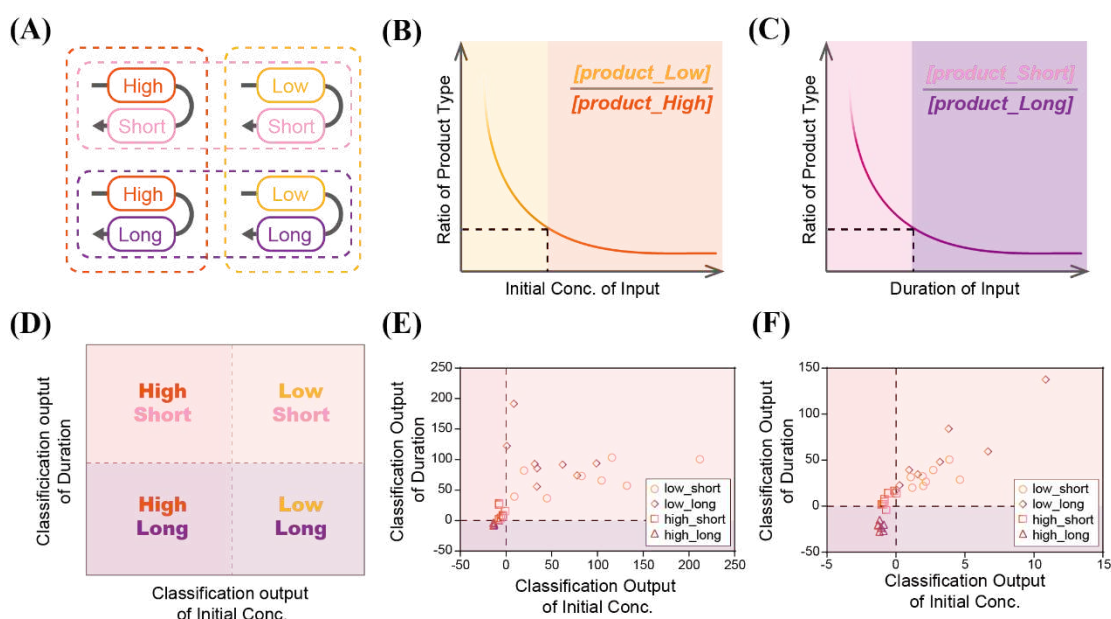


Figure 5. Single input, including concentration and duration information. (A) Test samples are grouped and combined for the further analysis. (B) The threshold is set by the critical initial concentration of Input. (C) The threshold is set by the critical time of duration of Input. (D) Both concentration and duration information can be divided into a 2D plot in which the classification output of initial concentration served as a horizontal axis, and the classification output of duration served as a vertical axis. Note that *Classification output* = *Ratio* – *Threshold*. (E) Scatter distribution of Input_A. (F) Scatter distribution of Input_B.

For precise prediction, we need to quantify if the predicted model is good enough. Therefore, we calculated the confusion matrixes of each module and discussed their advantages and limitation. Receiver-operating characteristic analysis of each module in the recording system are shown in Figure 6A, indicated that duration modules is less robust on classification compared to order and concentration modules. We found that the order module also works well (Accuracy $\approx$ 88%) in identifying the present order of A and B (Figure 6B), while concentration modules had superior performance (Accuracy = 100%) on classification both in A (Figure 6C) and B (Figure 6D). However, duration modules showed difficulties in classifying the “Low_Long” situation (Sensitivity = 50%, Accuracy = 75%) as discussed in the single-input section (Figure 6E and 6F).

To read out a molecular event, we assumed that all primers in the recording system would be used to generate recording products. Therefore, classification output could be calculated using the proportion of each recording product, following the hypothesis

described in the single-input section. We then calculated the total accuracy of a molecular event, and the results showed an accuracy of 68.75% in our 32 test samples (Supplementary Section 4.2.2).

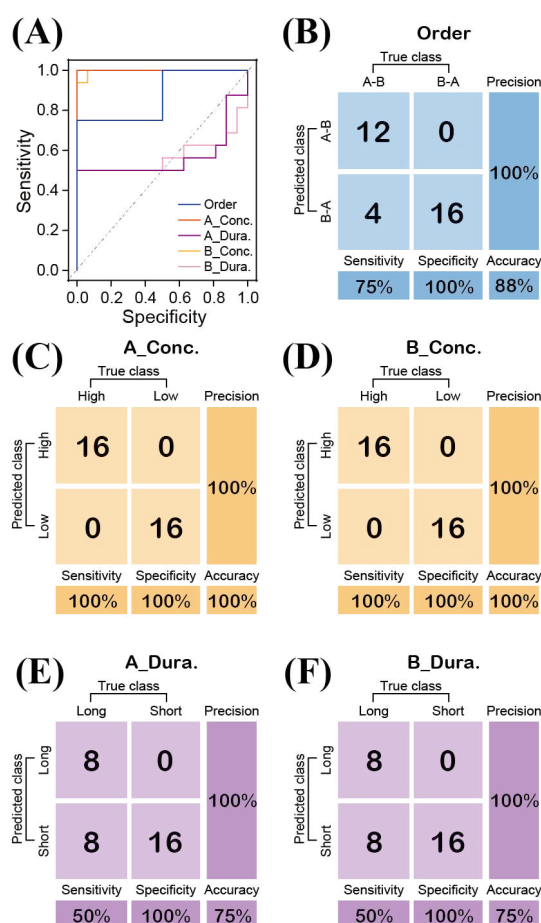


Figure 6. Classification model quantification. (A) Receiver-operating characteristic (ROC) curves for the order identification module, concentration classification modules, and duration classification modules. (B)-(F) Confusion matrixes for concentration classification, duration classification, and order identification.

Conclusions and Discussion

Here, we developed a one-pot molecular event recorder with simultaneous detecting and sequential recording. It contains three different information sensing modules: order, concentration, and duration. We first designed some DNA-based networks to capture the information of these parameters and then built several simulation models individually to predict event information through recording products. We tested the ability and stabilization of information transformation from Input to specific signals and then switched to the specific writer hairpin to record the information in each

functional module. We found that information sensing and transformation could be successfully achieved. Variants in these models were discussed, as well as their effect on recording products. We thus compared model-based prediction of the readouts and real situations in the test samples, suggesting that concentration and order modules showed high accuracy in their prediction, while duration modules still have limitations.

The restoration of information and its subsequent analysis were complicated by the appearance of 16 kinds of recording products, but in different proportions. Simulation modeling of each module showed monotonicity in the ratio of recording products, which we could use to restore the information of molecular events and work backward to a precise answer.

Our goal was the development of parallel functional modules to sense different sources of information. However, single-input data only work with one module. For example, the duration module would consume part of the Input to transform temporal information; however, this would affect concentration sensing by making it lower than true concentration. To reverse the errors, we lowered the working concentration in each same-level transducer in the duration module, as well as considering this situation in the concentration sensing model.

We demonstrated two different Inputs using DNA oligonucleotides. However, other types of target molecules, such as proteins or other small chemical molecules, can be chosen in further research. Momentary molecular events stored in DNA barcodes could support a reliable long-term analysis of significant molecules that perform inconsistently in biological activities. Examples are mRNA transcribed into complementary DNA (cDNA) by reverse transcription PCR (RT-PCR) ³³⁻³⁵, spatial transcriptomics in single cell ³⁶, cellular barcoding ³⁷. Such information storage and retrieval are important for cell fate identification, disease diagnosis, and treatment plan decisions. From our perspective, this simultaneous multi-parameter recording approach might sequester data of interest to multiple users in an efficient way.

Methods

DNA Synthesis and Purification. All DNA oligos were ordered from Genescript.

Purified DNA molecules were ordered purified with PAGE or HPLC. Oligos were suspended in 1× TAE Mg^{2+} buffer (Tris-acetate-EDTA buffer containing 12.5 mM Mg^{2+} , 0.04 M Tris Acetate, 1 mM EDTA, and 12.5 mM Mg Acetate, pH 8.3) buffer at 100 μM , and concentrations were measured using a Nanodrop and their extinction coefficients. All oligo mixtures listed in Supplementary Section 1.3 were diluted in 1× TAE Mg^{2+} to their working concentrations. Working solutions were stored at 4 °C, and stock solutions of DNA were stored at −20 °C. Oligo sequences for all experiments are listed in Supplementary Section 1.2.

PER Incubation. All PER experiments were incubated at 37 °C for the indicated times, usually with 1× ThermoPol buffer with supplemented magnesium (20 mM Tris-HCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM KCl, 12 mM MgSO_4 and 0.1% Triton X-100) and 0.8 units per μl of *Bst*. Large Fragment polymerase (purchased from New England Biolabs, M0275L) and 1 μM of the appropriate dHTPs (without dGTP, purchased from Diamond). Typically, 20 μl reactions were quenched by heat inactivation of the enzyme at 85 °C for 20 min. See Supplementary Section 2 for the reaction details of each experiment.

Gel Electrophoresis. All experiments used 15% TBE PAGE nondenaturing gels, which were run at 110 V for 110 min at 4 °C and scanned with the FAM and Cy5 channels. Gels were also stained with 4S GelRed if needed. Some experiments used different gel conditions.

Kinetics Experiments. Fluorescence measurements of the time delay module were performed using a Roche Light Cycler 480 (LC480) RT-qPCR system, and fluorescence kinetics data were collected following the list in Supplementary Section 2.2. Experiments were performed in wells of a 96-well plate with 20 μL reaction mixture per well at 37°C. Arbitrary fluorescence units were normalized to standard intensity by defining the minimum and maximum from the reaction sample without Input and delay signal with the same concentration of Timer-delay_signal in test samples.

Sequencing. All recording products were linked with UMI and group code in house (see Supplementary Section 2.6) and sequenced in Genewiz using next-generation sequencing.

Readout Data Analysis. For each module in this work, we recognized that the ratio

of the recording products all changes under monotonicity, i.e., neither decreasing nor increasing the initial concentration of the Input, the ratio of product_Low and product_High would increase or decrease (Figure S28A and S28B), thus enabling estimation of the initial concentration of Input by referencing the ratio. We grouped and combined the recording products and calculated by assuming that products of “low concentration” or “high concentration” are only tuned by the concentration module, while products with “short duration” or “long duration” are only affected by duration module (Supplementary Section 2.7). Based on this hypothesis, the output can be analyzed using the concentration module and duration module separately. At the end of an event recording, four types of products, including Low-Short, Low-Long, High-Short, and High-Long, would be produced in different proportions. Quantitative sequencing can help analyzing the proportion of each recording product through the frequency of each type of recording products. The classification output is defined as *product ratio – threshold*, so that the boundary of different group can be set as zero.

Acknowledgments

The authors thank Q. Lin and C. Deng for coding and data analysis, J. Wang for insightful conversations and comments on the manuscript. This work was financially supported by National Natural Science Foundation of China (

_____).

Author contributions

M.Z., D.H. and W.T. conceived and designed the study; M.Z. conducted most of the experiments and simulations; C.Y. and R.S. supported the optimization of simulations; M.Z., C.Z., J.W., Q.M., L.Y. and D.H. supported the optimization and analyzed data; M.Z. D.H., R.S. and W.T. wrote the manuscript. All authors reviewed the manuscript and approved the final version.

Competing interests

Q.M. and C.Z. are employees of Intellinosis Biotech Co., Ltd. with equity in the company.

Additional information

Supplementary information is available for this paper at publisher's website. Correspondence and requests for materials should be addressed to W.T., or D.H., or R.S.

References

- 1 Fahim Farzadfard, T. K. L. Emerging applications for DNA writers and molecular recorders. *Science* (2018).
- 2 Kuo, H.-C. *et al.* Weak gene–gene interaction facilitates the evolution of gene expression plasticity. *BMC Biology* **21**, 57, doi:10.1186/s12915-023-01558-6 (2023).
- 3 Hurst, L. D., Pál, C. & Lercher, M. J. The evolutionary dynamics of eukaryotic gene order. *Nature Reviews Genetics* **5**, 299-310, doi:10.1038/nrg1319 (2004).
- 4 Arakawa, S., Hatano, Y. & Katagiri, K. Differential expression of mRNA for Th1 and Th2 cytokine-associated transcription factors and suppressors of cytokine signalling in peripheral blood mononuclear cells of patients with atopic dermatitis. *Clinical and Experimental Immunology* **135**, 505-510, doi:10.1111/j.1365-2249.2004.02405.x (2004).
- 5 He, Y. & Tian, Z. NK cell education via nonclassical MHC and non-MHC ligands. *Cellular & Molecular Immunology* **14**, 321-330, doi:10.1038/cmi.2016.26 (2017).
- 6 Gong, Z. *et al.* Immunosuppressive reprogramming of neutrophils by lung mesenchymal cells promotes breast cancer metastasis. *Science Immunology* **8**, eadd5204, doi:10.1126/sciimmunol.add5204 (2023).
- 7 Xiao, Y. *et al.* Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation. *Cancer Cell* **39**, 423-437.e427, doi:10.1016/j.ccell.2020.12.012 (2021).
- 8 Ballesteros, I. *et al.* Co-option of Neutrophil Fates by Tissue Environments. *Cell* **183**, 1282-1297.e1218, doi:10.1016/j.cell.2020.10.003 (2020).
- 9 Weinberg, B. H. *et al.* Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature Biotechnology* **35**, 453-462, doi:10.1038/nbt.3805 (2017).
- 10 Roquet, N., Soleimany, A. P., Ferris, A. C., Aaronson, S. & Lu, T. K. Synthetic recombinase-based state machines in living cells. *Science* **353**, aad8559, doi:10.1126/science.aad8559 (2016).
- 11 Farzadfard, F. & Lu, T. K. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science* **346**, 1256272, doi:10.1126/science.1256272 (2014).
- 12 Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P. & Endy, D. Amplifying Genetic Logic Gates. *Science* **340**, 599-603, doi:10.1126/science.1232758 (2013).
- 13 Siuti, P., Yazbek, J. & Lu, T. K. Synthetic circuits integrating logic and memory in

- living cells. *Nature Biotechnology* **31**, 448-452, doi:10.1038/nbt.2510 (2013).
- 14 Perli, S. D., Cui, C. H. & Lu, T. K. Continuous genetic recording with self-targeting CRISPR-Cas in human cells. *Science* **353**, aag0511, doi:10.1126/science.aag0511 (2016).
- 15 Tang, W. & Liu, D. R. Rewritable multi-event analog recording in bacterial and mammalian cells. *Science* **360**, eaap8992, doi:10.1126/science.aap8992 (2018).
- 16 Kempton, H. R., Love, K. S., Guo, L. Y. & Qi, L. S. Scalable biological signal recording in mammalian cells using Cas12a base editors. *Nature Chemical Biology* **18**, 742-750, doi:10.1038/s41589-022-01034-2 (2022).
- 17 Choi, J. *et al.* A time-resolved, multi-symbol molecular recorder via sequential genome editing. *Nature* **608**, 98-107, doi:10.1038/s41586-022-04922-8 (2022).
- 18 Shipman, S. L., Nivala, J., Macklis, J. D. & Church, G. M. Molecular recordings by directed CRISPR spacer acquisition. *Science* **353**, aaf1175, doi:10.1126/science.aaf1175 (2016).
- 19 Sheth, R. U., Yim, S. S., Wu, F. L. & Wang, H. H. Multiplex recording of cellular events over time on CRISPR biological tape. *Science* **358**, 1457-1461, doi:10.1126/science.aao0958 (2017).
- 20 Schmidt, F., Cherepkova, M. Y. & Platt, R. J. Transcriptional recording by CRISPR spacer acquisition from RNA. *Nature* **562**, 380-385, doi:10.1038/s41586-018-0569-1 (2018).
- 21 Yim, S. S. *et al.* Robust direct digital-to-biological data storage in living cells. *Nature Chemical Biology* **17**, 246-253, doi:10.1038/s41589-020-00711-4 (2021).
- 22 Kishi, J. Y., Schaus, T. E., Gopalkrishnan, N., Xuan, F. & Yin, P. Programmable autonomous synthesis of single-stranded DNA. *Nat Chem* **10**, 155-164, doi:10.1038/nchem.2872 (2018).
- 23 Moerman, P. G., Gavrilov, M., Ha, T. & Schulman, R. Catalytic DNA Polymerization Can Be Expedited by Active Product Release. *Angew Chem Int Ed Engl* **61**, e202114581, doi:10.1002/anie.202114581 (2022).
- 24 Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585-1588, doi:10.1126/science.1132493 (2006).
- 25 Qian, L. & Winfree, E. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **332**, 1196, doi:10.1126/sciadv.ade0453 (2011).
- 26 Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368-372, doi:10.1038/nature10262 (2011).
- 27 Zhang, D. Y. & Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J Am Chem Soc* **131**, 17303-17314, doi:10.1021/ja906987s (2009).
- 28 Zhang, D. Y., Turberfield, A. J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318**, 1121-1125, doi:10.1126/science.1148532 (2007).
- 29 Aubert, N., Rondelez, Y., Fujii, T. & Hagiya, M. Enforcing logical delays in DNA computing systems. *Natural Computing* **13**, 559-572, doi:10.1007/s11047-014-9450-9 (2014).
- 30 Fern, J. *et al.* DNA Strand-Displacement Timer Circuits. *ACS Synthetic Biology* **6**, 190-193, doi:10.1021/acssynbio.6b00170 (2016).
- 31 Scalise, D. *et al.* Programming the Sequential Release of DNA. *ACS Synth Biol* **9**, 749-755, doi:10.1021/acssynbio.9b00398 (2020).
- 32 Chen, Y.-J. *et al.* Programmable chemical controllers made from DNA. *Nature Nanotechnology* **8**, 755-762, doi:10.1038/nnano.2013.189 (2013).
- 33 Zhang, C. *et al.* Cancer diagnosis with DNA molecular computation. *Nature Nanotechnology* **15**, 709-715, doi:10.1038/s41565-020-0699-0 (2020).
- 34 Ma, Q. *et al.* An automated DNA computing platform for rapid etiological diagnostics. *Science Advances* **8**, eade0453, doi:10.1126/sciadv.ade0453 (2022).
- 35 Zhang, C. *et al.* Logical Analysis of Multiple Single-Nucleotide-Polymorphisms with Programmable DNA Molecular Computation for Clinical Diagnostics.

- Angewandte Chemie International Edition* **61**, doi:10.1002/anie.202117658 (2022).
- 36 Crosetto, N., Bienko, M. & van Oudenaarden, A. Spatially resolved transcriptomics and beyond. *Nature Reviews Genetics* **16**, 57-66, doi:10.1038/nrg3832 (2015).
- 37 Kebschull, J. M. & Zador, A. M. Cellular barcoding: lineage tracing, screening and beyond. *Nature Methods* **15**, 871-879, doi:10.1038/s41592-018-0185-x (2018).