



**HAL**  
open science

## Machine Learning to Improve the Sensing of Biomolecules by Conical Track-Etched Nanopore

Nathan Meyer, Jean-Marc Janot, Mathilde Lepoitevin, Michael Smietana, Jean-Jacques Vasseur, Joan Torrent, Sebastien Balme

► **To cite this version:**

Nathan Meyer, Jean-Marc Janot, Mathilde Lepoitevin, Michael Smietana, Jean-Jacques Vasseur, et al.. Machine Learning to Improve the Sensing of Biomolecules by Conical Track-Etched Nanopore. Biosensors, 2020, 10 (10), pp.140. 10.3390/bios10100140 . hal-02997802

**HAL Id: hal-02997802**

**<https://hal.science/hal-02997802>**

Submitted on 10 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Article

# Machine Learning to Improve the Sensing of Biomolecules by Conical Track-Etched Nanopore

Nathan Meyer <sup>1,2</sup>, Jean-Marc Janot <sup>1</sup>, Mathilde Lepoitevin <sup>3</sup> , Michaël Smietana <sup>4</sup> ,  
Jean-Jacques Vasseur <sup>4</sup> , Joan Torrent <sup>2</sup> and Sébastien Balme <sup>1,\*</sup>

<sup>1</sup> Institut Européen des Membranes, UMR5635, UM, ENSCM, CNRS, 34095 Montpellier, France; nathan.meyer@umontpellier.fr (N.M.); jmjanot@univ-montp2.fr (J.-M.J.)

<sup>2</sup> Mécanismes Moléculaires dans les Démences Neurodégénératives, U1198, UM, EPHE, INSERM, 34095 Montpellier, France; joan.torrent@inserm.fr

<sup>3</sup> Institut des Matériaux Poreux de Paris UMR8004, CNRS, ENS, ESPCI, 75005 Paris, France; mathilde.lepoitevin@ens.fr

<sup>4</sup> Institut des Biomolécules Max Mousseron, Université de Montpellier, CNRS, ENSCM, 34095 Montpellier, France; michael.smietana@umontpellier.fr (M.S.); jean-jacques.vasseur@umontpellier.fr (J.-J.V.)

\* Correspondence: sebastien.balme@umontpellier.fr

Received: 7 September 2020; Accepted: 30 September 2020; Published: 5 October 2020



**Abstract:** Single nanopore is a powerful platform to detect, discriminate and identify biomacromolecules. Among the different devices, the conical nanopores obtained by the track-etched technique on a polymer film are stable and easy to functionalize. However, these advantages are hampered by their high aspect ratio that avoids the discrimination of similar samples. Using machine learning, we demonstrate an improved resolution so that it can identify short single- and double-stranded DNA (10- and 40-mers). We have characterized each current blockade event by the relative intensity, dwell time, surface area and both the right and left slope. We show an overlap of the relative current blockade amplitudes and dwell time distributions that prevents their identification. We define the different parameters that characterize the events as features and the type of DNA sample as the target. By applying support-vector machines to discriminate each sample, we show accuracy between 50% and 72% by using two features that distinctly classify the data points. Finally, we achieved an increased accuracy (up to 82%) when five features were implemented.

**Keywords:** nanopore; machine learning; DNA sensing

## 1. Introduction

For the past three decades, single nanopore technology have emerged as single-molecule sensors and offer many practical uses such as long read DNA sequencing [1,2]. This was achieved by engineering biological nanopores combined with biological machines to control the DNA translocation speed [3–7]. Beside sequencing, biological nanopores provide a nice platform to analyze the DNA substructure such as hairpin [8], the hybridization [9,10], zipping [11] or the interaction with protein [12]. At the beginning of the 2000s, the idea to mimic biological nanopores was demonstrated using different types of thin film. First, thin films of semiconductors (SiN) drilled by transmission electron microscopy or focused ion beam were used to provide nanopores with a low aspect ratio [13,14]. Next, polymer nanopore obtained by the track-etched technique provided a long high-aspect-ratio nanochannel [15,16]. More recently, 2D materials with reduced thickness down to a couple of angstroms, such as metal nitride or oxide, were developed to improve the noise and/or wettability of those low-aspect-ratio nanopore [17,18].

Regardless the type of artificial nanopore (solid-state or polymer), their performance in terms of precision is lower than the biological ones [19]. However, they offer pores of various sizes from a couple of nm up to hundreds of nm allowing the detection of folded proteins, protein assemblies and nanoparticles [20–22]. Artificial nanopores can be classified according to their aspect ratio. The 2D materials are the most promising to discriminate single nucleotides with potential applications in DNA sequencing [23,24]. The SiN and other nitride-based materials are the most used for single molecule sensing. They can be drilled by dielectric breakdown making this approach a low-cost technology [25,26]. Their aspect ratio allows discriminating the length of DNA as well as protein shape based on the amplitude of the current blockade but also the dwell time [27–33]. Solid-state nanopores also show interesting applications in the characterization of DNA knots and DNA-protein binding [34–36]. The track-etched nanopores have low resolution but their  $\mu\text{m}$  length scale allows increasing the dwell time making the polymer detection easier [37,38]. Even if track-etched nanopores are much less used than SiN membrane nanopores, several biomacromolecules were successfully detected such as proteins [37], DNA from 10 to 100 bp [39–42], hyaluronic acid [21] and amyloid [33,43]. In addition, they are mechanically robust, with expanded lifetimes up to several weeks [33]. This is particularly useful to investigate the kinetic of protein aggregation as well as their enzymatic degradation [20]. Furthermore, since the dwell time is enhanced in the track-etched nanopore [38], transient conductivity events can be easily detected without the use of a MHz amplifier. Nevertheless, their main limitation is their low resolution to discriminate small polymers. In order to improve their performance, their surfaces can be easily functionalized to tune their properties. Among them, the partial conversion of carboxylic acids into amine moieties [44], the deposition of  $\text{Al}_2\text{O}_3$  to tune the pore size [40] or the direct insertion of a biological nanopore [38]. Despite interesting results, the question of how the resolution of track-etched nanopore can be improved without chemical functionalization is still open. By tackling the problem of low resolution, the track-etched nanopore could offer a powerful platform to analyze the DNA size and structure. Indeed, compared to biological nanopore the tip diameter can be tuned to be sensitive to double-strand DNA or chain structure (i.e., knot, or hairpin). In this case, we could consider to determinate the ratio of different structures. Another advantage is the facility to modify the nanopore entrance to generate a specific interaction.

A way to increase resolution of nanopores is the use of machine learning algorithms [45,46], such as those used for DNA sequencing [47–50]. Moreover, combining the high resolution of biological nanopores and machine learning is a powerful tool to improve the nanopore resolution [51,52] allowing the identification of the protein domains [53], DNA base modification [54] and the C5 cytosine variant of DNA [55]. Here, we sought to apply classic algorithms of machine learning in the case of track-etched nanopores. Usually, the molecule detection by nanopore is characterized by two parameters: the relative current blockade and dwell time. The area of the events is also sometimes considered. There is compelling evidence using SiN nanopores that all these parameters are more or less correlated. We addressed the hypothesis that machine learning, in conjunction with the careful choice of multiple parameters that allow for characterizing the current blockade events, could improve the accuracy of DNA discrimination.

Here, we aim to demonstrate that classic algorithms of machine learning are powerful methods to data analysis of the nanopore sensing experiment. To do so, we have selected oligonucleotides with well-defined DNA sequences ( $A_{40}/T_{40}$ ,  $T_{40}$  and  $A_{10}/T_{10}$ ) as small macromolecules. Their detection was achieved through non-functionalized conical nanopore obtained by the track-etched method. The latter has a low resolution to discriminate small macromolecules. It is thus an ideal candidate to evaluate the benefits of the machine learning approach. From the parameter of the current blockade, we establish the correlation degree and then evaluate the accuracy of nanopores to discriminate the sample.

## 2. Material and Methods

### 2.1. Material

The A<sub>10</sub>, A<sub>40</sub>, T<sub>10</sub> and T<sub>40</sub> were obtained as previously reported [40]. Briefly, they were synthesized from commercially available phosphoramidite building blocks (Link Technologies Ltd., Bellshill, Scotland) in a 1 μmol scale using an ABI 381A DNA synthesizer by standard phosphoramidite chemistry. Then they were purified by RP-HPLC and characterized by MALDI-TOF MS.

### 2.2. Track-Etched Nanopore Design

Single conical nanopore was obtained by the track-etched method under dissymmetrical condition as previously reported [56]. Briefly, the single tracks were produced by Xe irradiation (8.98 MeV u<sup>-1</sup>) (GANIL, SME line, Caen, France) of polyethylene terephthalate (PET) film (thickness 13 μm, biaxial orientation ES30<sub>10</sub>61 Goodfellow). The tracks were activated by UV exposition 12 h per side (Fisher bioblock; VL215.MC, λ = 312 nm) before chemical etching process. The etching of conical nanopore was performed under dissymmetric condition (etchant solution 9 M NaOH and stop solution 1 M KCl 1 M of acetic acid) using the electrostopping method (1 V). After nanopore opening, the tip diameter ( $d_t$ ) of conical nanopores was calculated from the dependence of the conductance  $G$  (measured from −100 mV to 100 mV) with KCl concentration 1M, assuming bulk-like ionic conductivity inside the nanopores using Equation (1).

$$G = \frac{\kappa \pi d_t d_b}{4L} \quad (1)$$

where  $\kappa$  is the conductivity of the solution,  $L$  the nanopore length (13 μm) and  $d_b$  the diameter of the base side.  $d_b$  is calculated from the total etching time  $t$  using the relationship  $d_b = 2.5t$ . The factor 2.5 was determined in our laboratory using multipore membrane track. The pore dimensions used here are  $d_t = 3$  nm  $d_b = 200$  nm  $\alpha = 0.4^\circ$  (noted pore 1) and  $d_t = 4$  nm  $d_b = 350$  nm  $\alpha = 0.8^\circ$  (noted pore 2).

### 2.3. DNA Detection and Analysis

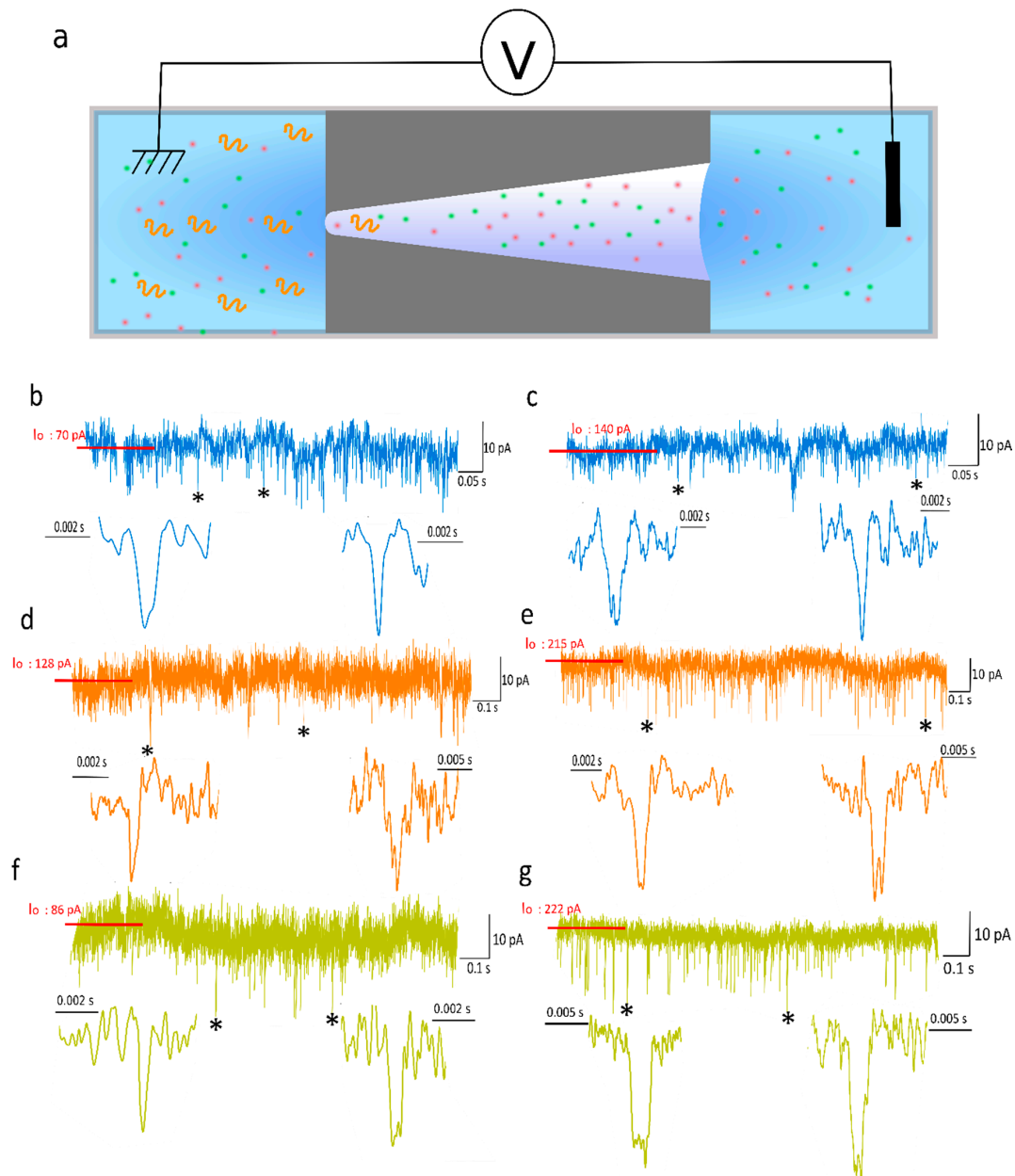
The DNA strands were detected using resistive pulse methods [57–59]. Briefly, the single conical nanopore was mounted between two Teflon chambers containing the same electrolyte solution (NaCl 3 M, EDTA 1 mM, PBS 50 mM, pH 7.2 or KCl 2 M, EDTA 1 mM, PBS 50 mM, pH 7.2). The current was measured by Ag/AgCl, 1 M KCl electrodes connected to the cell chambers by agar–agar bridges. The working electrode and ground electrode were located in the trans-chamber (base side of the nanopore) and in the cis chamber (tip side of the nanopore), respectively. Electrical measurement was performed using a patch-clamp amplifier (EPC10 HEKA electronics, Lambrecht, Germany).

The polynucleotide samples were added on the cis chamber (tip side of nanopore) to reach a final concentration of 10 nM. Positive bias (250 mV or 500 mV) was then applied to the trans-chamber. Ion current was recorded at a sampling frequency of 100 kHz (for T<sub>40</sub> and A<sub>40</sub>/T<sub>40</sub>) or 200 kHz (for A<sub>10</sub>/T<sub>10</sub>). A Bessel filter at 10 kHz is used. Those experiments were repeated at least 10 times in 8 successive days for each nanopore. The data analysis was performed using a custom-made LabView software with Butterworth filter of 2.5 kHz, 2 orders. The base line fluctuation was corrected using a Savitzky–Golay filter of 2400 side points, 1 order. The detection event was performed using a threshold of  $3\sigma$  ( $\sigma$  where is the standard deviation of the signal). Each event was characterized by the relative current blockade ( $\Delta I/I_0$ ), the dwell time ( $\Delta t$ ), the area ( $AUC$ ), the right ( $RS$ ) and left slopes ( $LS$ ). The parameters of the current blockade were analyzed using Matlab and the toolbox “statistical and learning machine”.

## 3. Results and Discussion

The experimental detection of all DNA samples A<sub>10</sub>/T<sub>10</sub>, A<sub>40</sub>/T<sub>40</sub> and T<sub>40</sub> were performed from the tip side to the base side under two different electrolyte conditions (NaCl 3 M, EDTA 1 mM, PBS 50 mM,

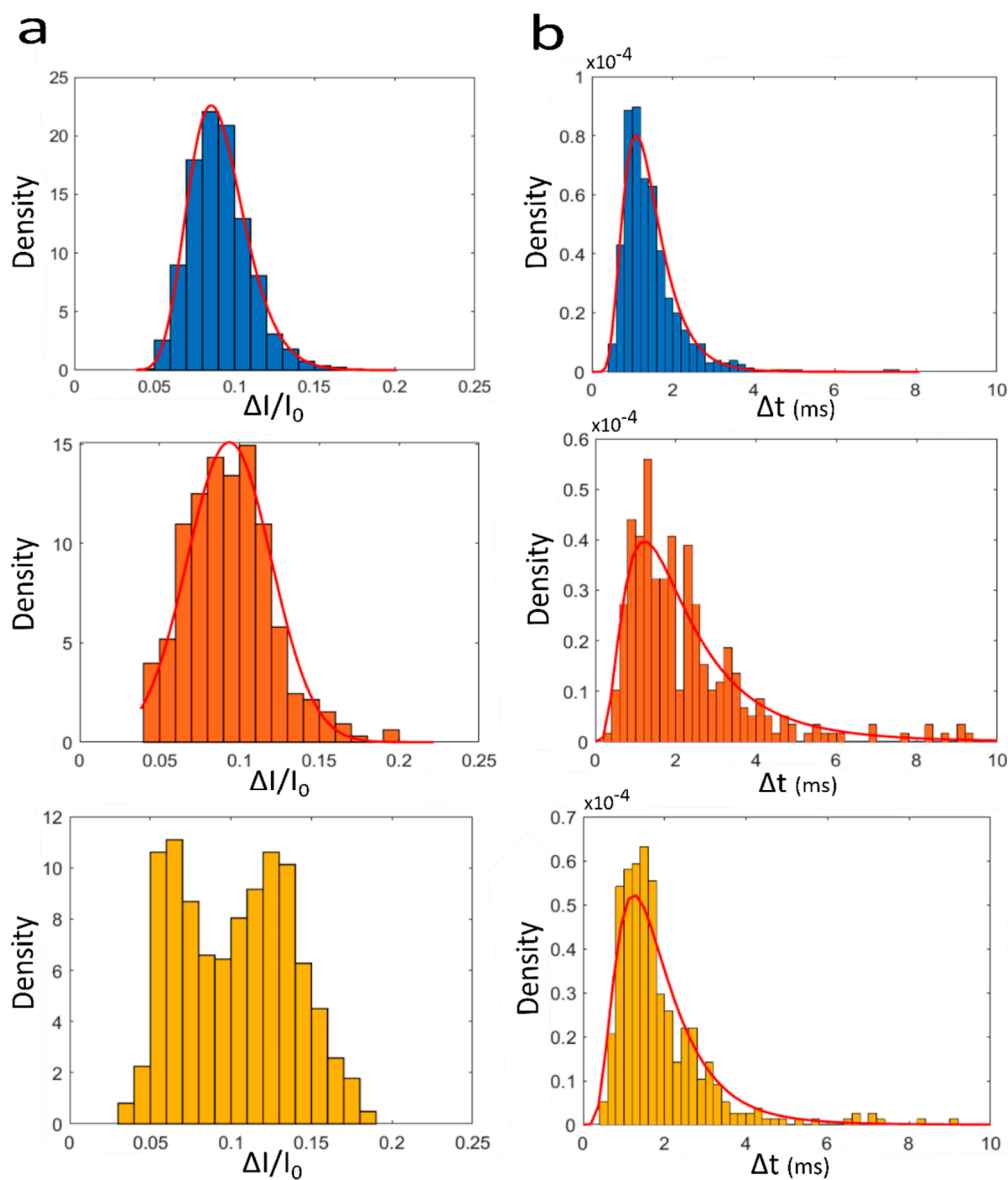
pH 7.2 or KCl 2 M, EDTA 1 mM, PBS 50 mM, pH 7.2) (Figure 1a). Figure 1b–g shows examples of current traces recorded at 250 mV and 500 mV for all samples. From the current traces, the events related to the DNA translocation through the nanopore were detected. These current blockades were usually described by the relative current blockade ( $\Delta I/I_0$ ), which is the ratio between the amplitude of the current blockade and the base line current, and the dwell time. These two parameters were first extracted to characterize all the events recorded during our experiments.



**Figure 1.** (a) Sketch of DNA sensing using conical nanopore. Examples of current traces and zooms of current blockade recorded for  $A_{10}/T_{10}$  at (b) 250 mV and (c) 500 mV,  $A_{40}/T_{40}$  at (d) 250 mV and (e) 500 mV,  $T_{40}$  at (f) 250 mV and (g) 500 mV. The symbol \* identifies the examples of current blockade zoom below each trace. The current traces were obtained using pore 1  $d_t = 3$  nm,  $d_b = 200$  nm.

In Figure 2 are reported the distribution histograms of  $\Delta I/I_0$  obtained for the  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  and  $T_{40}$  at a voltage of 500 mV using pore 1 under KCl 2 M. These distributions are centered at similar values for the three samples: 0.085, 0.093 and 0.066 (another center of distribution is observed at 0.128) for  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  and  $T_{40}$ , respectively. The  $\Delta I/I_0$  distributions of the current blockade recorded

at 250 mV are centered on 0.15, 0.09 and 0.12 for  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  and  $T_{40}$ , respectively (Figure S1). These values slightly increase for pore 2 (under NaCl 3 M) recorded at 250 mV. The distributions are centered on 0.27, 0.25 and 0.18 for  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  and  $T_{40}$ , respectively (Figure S2). Similar observation can be made for the experiments performed at 500 mV where the centers of distribution are 0.31, 0.28, and 0.36 for  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  and  $T_{40}$ , respectively (Figure S3).



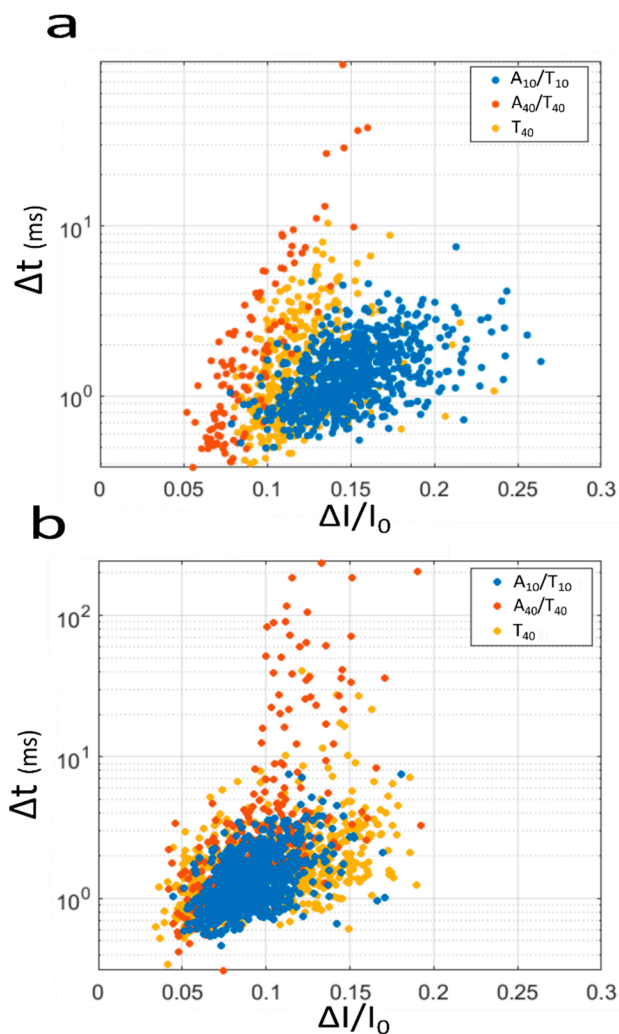
**Figure 2.** Distribution histograms of (a) the amplitude of the relative current blockade ( $\Delta I/I_0$ ) and (b) the dwell time ( $\Delta t$ ) for the  $A_{10}/T_{10}$  (blue),  $A_{40}/T_{40}$ : (orange),  $T_{40}$  (yellow). The events were recorded at 500 mV, the number of events  $n = 793, 332$  and  $634$  for  $A_{10}/T_{10}, A_{40}/T_{40}$  and  $T_{40}$ , respectively. The results were obtained using pore 1  $d_t = 3$  nm,  $d_b = 200$  nm. The density ( $d_i$ ) is the frequency ( $f_i$ ) of event relative to the sample size ( $n$ ) and the bin width ( $w_i$ )  $d_i = \frac{f_i}{nw_i}$  where  $w_i = 0.01$  and  $200$  for  $\Delta I/I_0$  and  $\Delta t$ , respectively.

The distribution of  $\Delta t$  for the three samples at 500 mV recorded for pore 1 (and 2) are reported in Figure 2b and Figure S2b. We observe that the distributions are centered close to the same value: 1.07 ms (1.48 ms), 0.95 ms (1.24 ms), and 1.29 ms (1.23 ms) for  $A_{10}/T_{10}, A_{40}/T_{40}$  and  $T_{40}$ , respectively. We notice that the time scale (about 1 ms) is in the same range as the one reported for DNA 50 bp [41]. Under 250 mV, the  $\Delta t$  values do not significantly decrease (Figures S2b and S4b). Indeed, the centers of



distribution for pore 1 (and 2) are found to be 1.07 ms (1.39 ms), 0.82 ms (1.25 ms) and 0.87 ms (1.30 ms) for  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  and  $T_{40}$ .

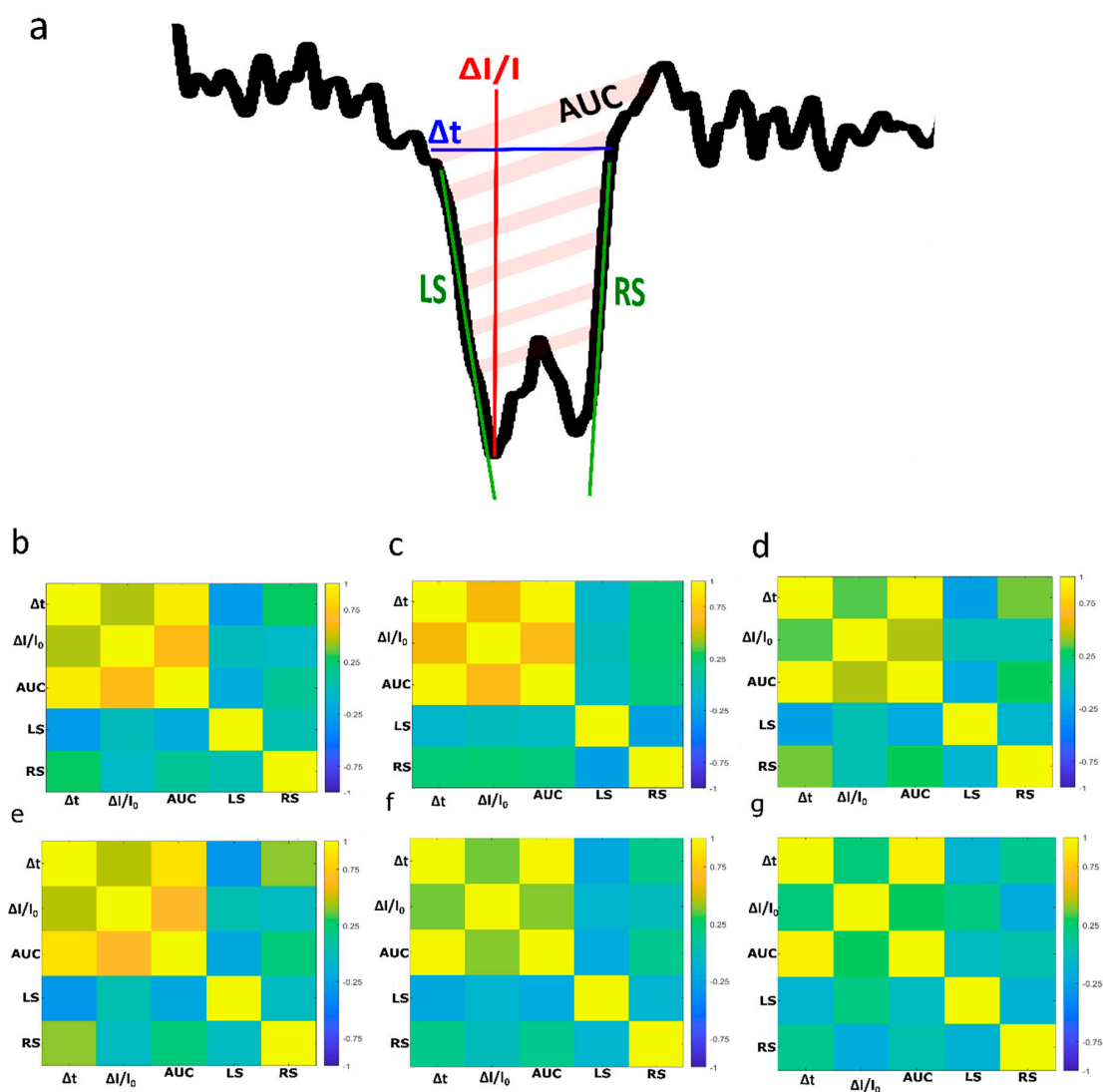
Usually, the  $\Delta I/I_0$  and  $\Delta t$  are the main parameters to discriminate the sample analyzed by nanopore sensors. Here, we observe a large overlap between the different distributions (Figure 2 and Figures S1–S3) preventing the sample discrimination. The results also indicate that there is no preferential voltage or pore to discriminate them with only one parameter. In Figure 3 are reported two event maps representing the  $\Delta I/I_0$  vs.  $\Delta t$  of translocation events for the three samples at two different voltages: 250 mV (Figure 3a) and 500 mV (Figure 3b) for pore 1. We observe that the cloud of events overlaps due to the similar distribution of  $\Delta I/I_0$  and  $\Delta t$ . This overlap makes impossible the discrimination of the DNA samples by a simple clustering analysis. The same trend is observed for pore 2 for the two same voltages (250 and 500 mV, see Figure S4). These observations are not surprising due to the low resolution of track-etched nanopore.



**Figure 3.** Scatter plot representing the  $\Delta t$  versus the  $I/I_0$  for the  $A_{10}/T_{10}$  (blue),  $A_{40}/T_{40}$  (orange),  $T_{40}$  (yellow) for a voltage of (a) 250 mV and (b) 500 mV. The results were obtained using the pore  $d_t = 3$  nm,  $d_b = 200$  nm. The number of events recorded at 250 mV  $n = 703$ , 116 and 382 and at 500 mV  $n = 793$ , 332 and 640 for  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  and  $T_{40}$ , respectively.

To go further, we attempted to define each current blockade with additional parameters (Figure 4a). First, we considered the surface area of the event (AUC) because it takes into account the eventual current fluctuation during the DNA translocation. We could expect that this parameter is strongly correlated to the  $\Delta t$  and the  $\Delta I/I_0$ . The conical shape of the nanopore can generate a dissymmetrical

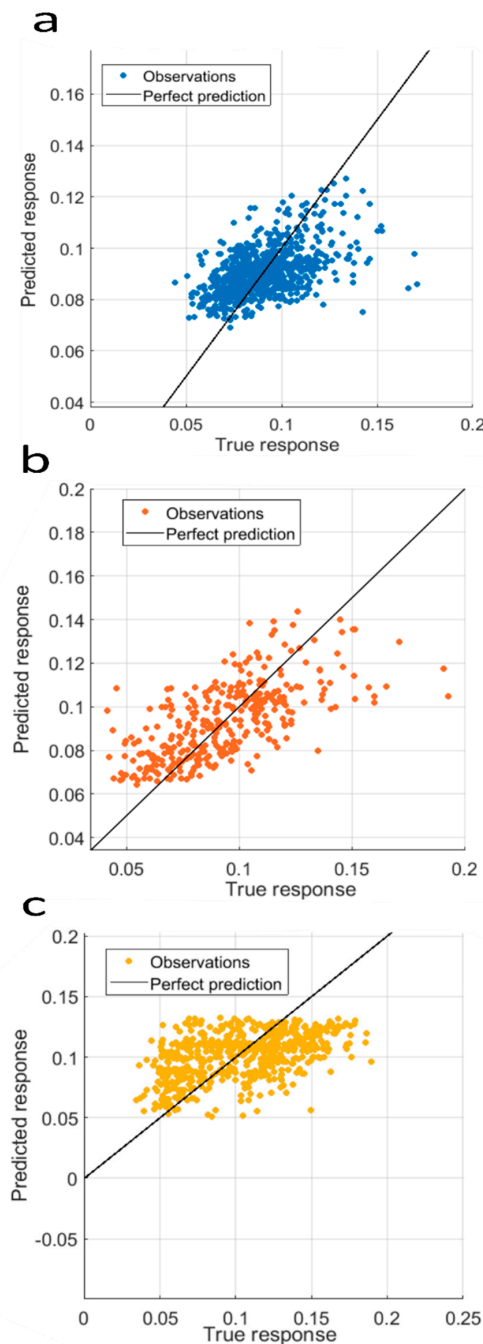
shape of current blockade events. In that case, the event's right and left slopes (noted RS and LS, respectively) are expected to be different as previously reported in the case of spherical object [60,61]. Now, we evaluate the correlation degree of these five parameters ( $\Delta t$ ,  $\Delta I/I_0$ , AUC, RS and LR). Usually, a positive correlation between the  $\Delta t$  and the  $\Delta I/I_0$  can be observed if the length of the pore is close to that of the analyte. Indeed, this correlation has been reported for protein detection using SiN nanopore [62]. Conversely, in the case of long DNA strands, the amplitudes of the relative current blockade are not correlated with the dwell time since the nanopore is filled with the polymer strand [63]. Here, we report the correlation heat maps of various parameters for the three samples at a voltage of 250 mV (top line) and 500 mV (bottom line) for pore 1 (Figure 4b) and pore 2 (Figure S5). For all samples and regardless of the pore or the applied voltage, we can observe a strong correlation between  $\Delta t$  and the surface area ( $\sim 0.90$  in mean). Conversely, the correlation between the  $\Delta t$  and  $\Delta I/I_0$  is low ( $< 0.75$ ). This low correlation degree is also observed between the surface and the  $\Delta I/I_0$ . This could be explained by the current fluctuation during the blockade due to the DNA motion inside the pore. Interestingly, the right and the left slopes do not appear to be correlated to each other ( $\sim -0.20$  in mean) nor with other parameters.



**Figure 4.** (a) Zoom of a translocation event and the representation of variables characterizing events. (b–g) Heat map representing the correlation between the variables characterizing events obtained under 250 mV for (b) A<sub>10</sub>/T<sub>10</sub>, (c) A<sub>40</sub>/T<sub>40</sub>, (d) T<sub>40</sub> and at 500 mV for (e) A<sub>10</sub>/T<sub>10</sub>, (f) A<sub>40</sub>/T<sub>40</sub>, (g) T<sub>40</sub>. The results were obtained using the pore  $d_t = 3$  nm,  $d_b = 200$  nm.



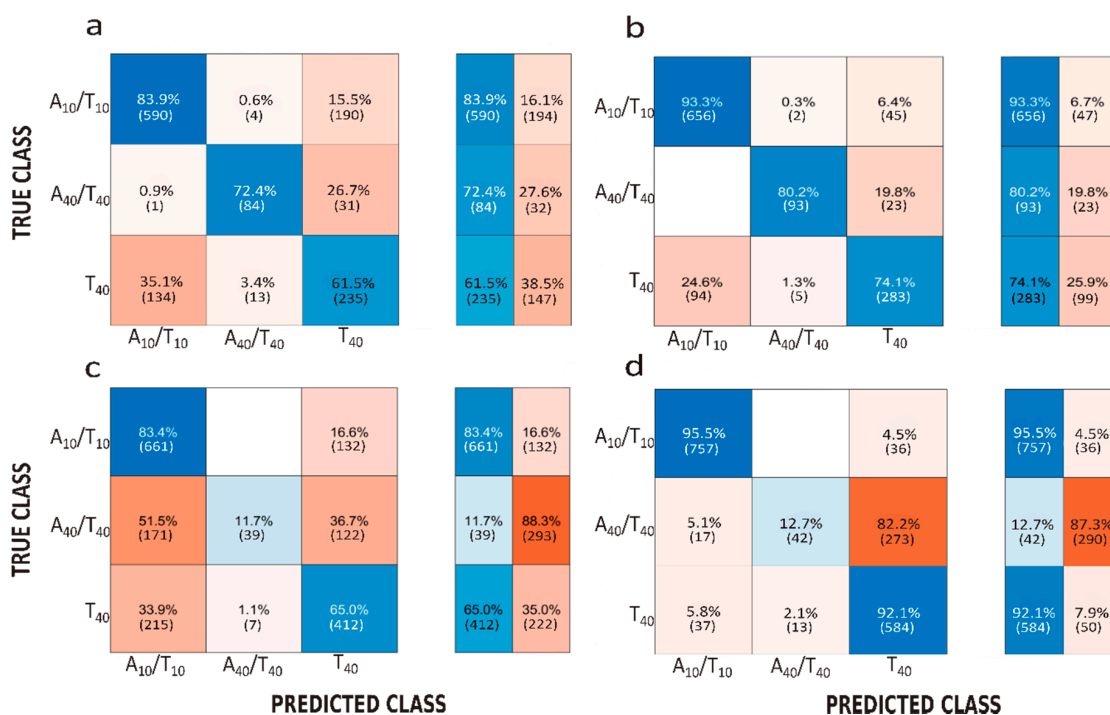
We then attempted to improve sample discrimination using machine learning algorithms. The simplest model involves establishing a linear correlation between two parameters. First, we examined whether  $\Delta I/I_0$  and the  $\Delta t$  are correlated. In Figure 5 is reported the linear regression analysis performed with  $\Delta I/I_0$  as response variable and  $\Delta t$  as predictor at a voltage of 500 mV for pore 1. We can observe a low correlation between these two parameters according to a  $R^2$  about 0.25 in mean. The same analysis for the event recorded at 250 mV and with pore 2 (Figure S6) also provides a low  $R^2$  value (about 0.21). This is in good agreement with the heat map and confirms the non-linearity between the  $\Delta I/I_0$  and the  $\Delta t$ .



**Figure 5.** Linear regression performed with  $I/I_0$  (response variable) and  $\Delta t$  (predictor) for (a)  $A_{10}/T_{10}$  ( $R^2 = 0.23$ ;  $n = 793$ ), (b)  $A_{40}/T_{40}$  ( $R^2 = 0.43$ ;  $n = 332$ ) (c)  $T_{40}$  ( $R^2 = 0.23$ ;  $n = 634$ ). The results were obtained using the pore  $d_t = 3$  nm,  $d_b = 200$  nm, voltage of 500 mV.

The support vector machine is a class of machine learning algorithm used to solve classification problems. The data training involves finding a way to separate the different samples by using the different parameters that characterize the events. In our case, we have defined the sample  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  or  $T_{40}$  as the target of the algorithm. We have defined as the features the different parameters that characterize the events ( $\Delta I/I_0$ ,  $\Delta t$ , AUC, RS and RL). As previously mentioned, the  $\Delta I/I_0$ ,  $\Delta t$  are the most commonly used. Thus, we trained the algorithm with these two features and then with five features in order to demonstrate that the added parameters will help to improve the discrimination and to classify the different samples.

In Figure 6 is reported the confusion matrix for pore 1. Using two features ( $I/I_0$  and  $\Delta t$ ), the accuracy is 72.6% and 53% for the event recorded at 250 mV and 500 mV, respectively. First, we observe that using machine learning the accuracy is better at 250 mV making this voltage more relevant to discriminate samples. Ignoring the voltage, the best predictions of event parameters were found for the  $A_{10}/T_{10}$ . As expected, the use of five features allows for improving the accuracy up to 82.5% and 66.76% at 250 mV and 500 mV, respectively. This improvement of the accuracy is also observed for pore 2 (Figure S7). This weak difference in precision between the two pores is likely due to the different geometries. The best results were obtained with the smaller nanopores. Using five features, the ratio of the true positive is higher for the  $A_{10}/T_{10}$  and  $T_{40}$  than for the  $A_{40}/T_{40}$ . This is observed for all experiments except for pore 2 at 250 mV. However, this could be explained by a low number of events ( $n = 80$ ).



**Figure 6.** Confusion matrices representing the accuracy of classification with the support vector machine approach obtained under  $V = 250$  mV using (a) 2 features ( $\Delta I/I_0$ ,  $\Delta t$ ) and (b) 5 features ( $\Delta I/I_0$ ,  $\Delta t$ , AUC, LS, RS) and under 500 mV using (c) 2 features ( $\Delta I/I_0$ ,  $\Delta t$ ) and (d) 5 features ( $\Delta I/I_0$ ,  $\Delta t$ , AUC, LS, RS). The results were obtained using pore 1  $d_t = 3$  nm,  $d_b = 200$  nm.

#### 4. Conclusions

In summary, three DNA samples named  $A_{10}/T_{10}$ ,  $A_{40}/T_{40}$  or  $T_{40}$  were detected using conical nanopore with a tip diameter of about 3 nm. The classical parameters used to characterize the event ( $I/I_0$  and  $\Delta t$ ) do not allow one to discriminate the samples due to a large overlap of their distributions. In addition, the linear regression analysis shows no correlation between these two

parameters. Using support vector machines, the different samples were discriminated with accuracy between 50% and 72.6%. The events were then characterized by five parameters that are not correlated to each other except for  $\Delta t$  and the surface area. The introduction of three additional parameters as features (AUC, LS, RS) in the support vector machine algorithm showed 10% improvement of accuracy, which increased to 82.5% for the smallest nanopore at 250 mV. Among the three samples, the best prediction of event parameters was found for the  $A_{10}/T_{10}$ .

More generally, our work was motivated to propose a solution to improve the resolution of conical track-etched nanopore. The combination of additional parameters and support vector machine algorithms was found to be a relevant solution to reach this goal. We could expect that such analysis methodology will be applied for single molecule sensing using track-etched nanopores, especially in fields where these nanopores could bring fundamental insights, such as in amyloid detection and characterization.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2079-6374/10/10/140/s1>. Figure S1: Distribution histograms of  $\Delta I/I_0$  and  $\Delta t$  recorded with pore 1 at 250 mV, Figure S2: Distribution histograms of  $\Delta I/I_0$  and  $\Delta t$  recorded with pore 2 at 500 mV, Figure S3: Distribution histograms of  $\Delta I/I_0$  and  $\Delta t$  recorded with pore 2 at 250 mV, Figure S4: Scatter plot representing the  $\Delta t$  versus the  $\Delta I/I_0$  recorded with the pore 2, Figure S5: Heat map representing the correlation between the variables characterizing events obtained under 250 mV. Figure S6: Linear regression performed with  $\Delta I/I_0$  (response variable) and  $\Delta t$  (predictor), Figure S7: Confusion matrix representing the accuracy of classification with the support vector machine approach obtained under  $V = 250$  mV.

**Author Contributions:** Conceptualization, S.B.; software, J.-M.J.; validation, S.B., and M.L.; formal analysis, N.M.; investigation, N.M., M.L., M.S., S.B.; resources, J.-J.V., S.M.; writing—original draft preparation, N.M. and S.B.; writing—review and editing, M.L., M.S., J.-J.V., J.T.; visualization, N.M.; supervision, S.B., J.T.; funding acquisition, S.B. and J.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Agence Nationale de la Recherche (ANR-19-CE42-0006, NanoOligo). Single tracks have been produced at GANIL (Caen, France) in the framework of an EMIR project.

**Acknowledgments:** The authors acknowledge E. Balanzat (CIMAP/Caen) for assistance during swift heavy ion irradiation.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Deamer, D.; Akeson, M.; Branton, D. Three decades of nanopore sequencing. *Nat. Biotechnol.* **2016**, *34*, 518–524. [[CrossRef](#)]
2. Kasianowicz, J.J.; Brandin, E.; Branton, D.; Deamer, D.W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13770–13773. [[CrossRef](#)]
3. Manrao, E.A.; Derrington, I.M.; Laszlo, A.H.; Langford, K.W.; Hopper, M.K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J.H. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* **2012**, *30*, 349–353. [[CrossRef](#)] [[PubMed](#)]
4. Wloka, C.; Mutter, N.L.; Soskine, M.; Maglia, G. Alpha-Helical Fragaceatoxin C Nanopore Engineered for Double-Stranded and Single-Stranded Nucleic Acid Analysis. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 12494–12498. [[CrossRef](#)] [[PubMed](#)]
5. Van der Verren, S.E.; van Gerven, N.; Jonckheere, W.; Hambley, R.; Singh, P.; Kilgour, J.; Jordan, M.; Wallace, E.J.; Jayasinghe, L.; Remaut, H. A dual-constriction biological nanopore resolves homonucleotide sequences with high fidelity. *Nat. Biotechnol.* **2020**. [[CrossRef](#)]
6. Cressiot, B.; Bacri, L.; Pelta, J. The Promise of Nanopore Technology: Advances in the Discrimination of Protein Sequences and Chemical Modifications. *Small Methods* **2020**, 2000090. [[CrossRef](#)]
7. Asandei, A.; Di Muccio, G.; Schiopu, I.; Mereuta, L.; Dragomir, I.S.; Chinappi, M.; Luchian, T. Nanopore-Based Protein Sequencing Using Biopores: Current Achievements and Open Challenges. *Small Methods* **2020**, 1900595. [[CrossRef](#)]
8. Ding, Y.; Fleming, A.M.; White, H.S.; Burrows, C.J. Internal vs fishhook hairpin DNA: Unzipping locations and mechanisms in the  $\alpha$ -hemolysin nanopore. *J. Phys. Chem. B* **2014**, *118*, 12873–12882. [[CrossRef](#)]

9. Mereuta, L.; Asandei, A.; Schiopu, I.; Park, Y.; Luchian, T. Nanopore-Assisted, Sequence-Specific Detection, and Single-Molecule Hybridization Analysis of Short, Single-Stranded DNAs. *Anal. Chem.* **2019**, *91*, 8630–8637. [[CrossRef](#)]
10. Ciuca, A.; Asandei, A.; Schiopu, I.; Apetrei, A.; Mereuta, L.; Seo, C.H.; Park, Y.; Luchian, T. Single-Molecule, Real-Time Dissecting of Peptide Nucleic Acid–DNA Duplexes with a Protein Nanopore Tweezer. *Anal. Chem.* **2018**, *90*, 7682–7690. [[CrossRef](#)]
11. Perera, R.T.; Fleming, A.M.; Peterson, A.M.; Heemstra, J.M.; Burrows, C.J.; White, H.S. Unzipping of A-Form DNA–RNA, A-Form DNA–PNA, and B-Form DNA–DNA in the  $\alpha$ -Hemolysin Nanopore. *Biophys. J.* **2016**, *110*, 306–314. [[CrossRef](#)] [[PubMed](#)]
12. Hornblower, B.; Coombs, A.; Whitaker, R.D.; Kolomeisky, A.; Picone, S.J.; Meller, A.; Akeson, M. Single-molecule analysis of DNA–protein complexes using nanopores. *Nat. Methods* **2007**, *4*, 315–317. [[CrossRef](#)] [[PubMed](#)]
13. Li, J.; Stein, D.; McMullan, C.; Branton, D.; Aziz, M.J.; Golovchenko, J.A. Ion-beam sculpting at nanometre length scales. *Nature* **2001**, *412*, 166–169. [[CrossRef](#)]
14. Li, J.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J.A. DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mater.* **2003**, *2*, 611–615. [[CrossRef](#)]
15. Harrell, C.C.; Lee, S.B.; Martin, C.R. Synthetic single-nanopore and nanotube membranes. *Anal. Chem.* **2003**, *75*, 6861–6867. [[CrossRef](#)] [[PubMed](#)]
16. Heins, E.A.; Siwy, Z.S.; Baker, L.A.; Martin, C.R. Detecting single porphyrin molecules in a conically shaped synthetic nanopore. *Nano Lett.* **2005**, *5*, 1824–1829. [[CrossRef](#)]
17. Liu, K.; Feng, J.; Kis, A.; Radenovic, A. Atomically thin molybdenum disulfide nanopores with high sensitivity for DNA translocation. *ACS Nano* **2014**, *8*, 2504–2511. [[CrossRef](#)]
18. Schneider, G.F.; Kowalczyk, S.W.; Calado, V.E.; Pandraud, G.; Zandbergen, H.W.; Vandersypen, L.M.K.; Dekker, C. DNA translocation through graphene nanopores. *Nano Lett.* **2010**, *10*, 3163–3167. [[CrossRef](#)]
19. Hall, A.R.; Scott, A.; Rotem, D.; Mehta, K.K.; Bayley, H.; Dekker, C. Hybrid pore formation by directed insertion of  $\alpha$ -haemolysin into solid-state nanopores. *Nat. Nanotechnol.* **2010**, *5*, 874–877. [[CrossRef](#)]
20. Giamblanco, N.; Janot, J.-M.; Gubbiotti, A.; Chinappi, M.; Balme, S. Characterization of Food Amyloid Protein Digestion by Conical Nanopore. *Small Methods* **2020**, 1900703. [[CrossRef](#)]
21. Ma, T.; Janot, J.-M.; Balme, S. Track-Etched Nanopore/Membrane: From Fundamental to Applications. *Small Methods* **2020**, 2000366. [[CrossRef](#)]
22. Lepoitevin, M.; Ma, T.; Bechelany, M.; Janot, J.-M.; Balme, S. Functionalization of single solid state nanopores to mimic biological ion channels: A review. *Adv. Colloid Interface Sci.* **2017**, *250*, 195–213. [[CrossRef](#)]
23. Feng, J.; Liu, K.; Bulushev, R.D.; Khlybov, S.; Dumcenco, D.; Kis, A.; Radenovic, A. Identification of single nucleotides in MoS<sub>2</sub> nanopores. *Nat. Nanotechnol.* **2015**, *10*, 1070–1076. [[CrossRef](#)] [[PubMed](#)]
24. Thakur, M.; Macha, M.; Chernev, A.; Graf, M.; Lihter, M.; Deen, J.; Tripathi, M.; Kis, A.; Radenovic, A. Wafer-Scale Fabrication of Nanopore Devices for Single-Molecule DNA Biosensing using MoS<sub>2</sub>. *Small Methods* **2020**, 2000072. [[CrossRef](#)]
25. Kwok, H.; Briggs, K.; Tabard-Cossa, V. Nanopore fabrication by controlled dielectric breakdown. *PLoS ONE* **2014**, *9*, e92880. [[CrossRef](#)] [[PubMed](#)]
26. Briggs, K.; Kwok, H.; Tabard-Cossa, V. Automated fabrication of 2-nm solid-state nanopores for nucleic acid analysis. *Small* **2014**, *10*, 2077–2086. [[CrossRef](#)]
27. Balme, S.; Coulon, P.E.; Lepoitevin, M.; Charlot, B.; Yandrapalli, N.; Favard, C.; Muriaux, D.; Bechelany, M.; Janot, J.-M. Influence of Adsorption on Proteins and Amyloid Detection by Silicon Nitride Nanopore. *Langmuir* **2016**, *32*, 8916–8925. [[CrossRef](#)] [[PubMed](#)]
28. Kowalczyk, S.W.; Hall, A.R.; Dekker, C. Detection of local protein structures along DNA using solid-state nanopores. *Nano Lett.* **2010**, *10*, 324–328. [[CrossRef](#)]
29. Larkin, J.; Henley, R.; Bell, D.C.; Cohen-Karni, T.; Rosenstein, J.K.; Wanunu, M. Slow DNA transport through nanopores in hafnium oxide membranes. *ACS Nano* **2013**, *7*, 10121–10128. [[CrossRef](#)]
30. Venta, K.; Shemer, G.; Puster, M.; Rodríguez-Manzo, J.A.; Balan, A.; Rosenstein, J.K.; Shepard, K.; Drndić, M. Differentiation of short, single-stranded DNA homopolymers in solid-state nanopores. *ACS Nano* **2013**, *7*, 4629–4636. [[CrossRef](#)]
31. Wei, R.; Gatterdam, V.; Wieneke, R.; Tampé, R.; Rant, U. Stochastic sensing of proteins with receptor-modified solid-state nanopores. *Nat. Nanotechnol.* **2012**, *7*, 257–263. [[CrossRef](#)] [[PubMed](#)]

32. Smeets, R.M.M.; Keyser, U.F.; Krapf, D.; Wu, M.-Y.; Dekker, N.H.; Dekker, C. Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett.* **2006**, *6*, 89–95. [[CrossRef](#)] [[PubMed](#)]
33. Giambianco, N.; Coglitore, D.; Janot, J.-M.; Coulon, P.E.; Charlot, B.; Balme, S. Detection of protein aggregate morphology through single antifouling nanopore. *Sens. Actuators B Chem.* **2018**, *260*, 736–745. [[CrossRef](#)]
34. Plesa, C.; Verschuere, D.; Pud, S.; van der Torre, J.; Ruitenber, J.W.; Witteveen, M.J.; Jonsson, M.P.; Grosberg, A.Y.; Rabin, Y.; Dekker, C. Direct observation of DNA knots using a solid-state nanopore. *Nat. Nanotechnol.* **2016**, *11*, 1093–1097. [[CrossRef](#)] [[PubMed](#)]
35. Plesa, C.; Ruitenber, J.W.; Witteveen, M.J.; Dekker, C. Detection of Individual Proteins Bound along DNA Using Solid-State Nanopores. *Nano Lett.* **2015**, *15*, 3153–3158. [[CrossRef](#)] [[PubMed](#)]
36. Kumar Sharma, R.; Agrawal, I.; Dai, L.; Doyle, P.S.; Garaj, S. Complex DNA knots detected with a nanopore sensor. *Nat. Commun.* **2019**, *10*, 4473. [[CrossRef](#)]
37. Mara, A.; Siwy, Z.; Trautmann, C.; Wan, J.; Kamme, F. An Asymmetric Polymer Nanopore for Single Molecule Detection. *Nano Lett.* **2004**, *4*, 497–501. [[CrossRef](#)]
38. Cabello-Aguilar, S.; Balme, S.; Chaaya, A.A.; Bechelany, M.; Balanzat, E.; Janot, J.-M.; Pochat-Bohatier, C.; Miele, P.; Dejardin, P. Slow translocation of polynucleotides and their discrimination by  $\alpha$ -hemolysin inside a single track-etched nanopore designed by atomic layer deposition. *Nanoscale* **2013**, *5*, 9582–9586. [[CrossRef](#)]
39. Harrell, C.C.; Choi, Y.; Horne, L.P.; Baker, L.A.; Siwy, Z.S.; Martin, C.R. Resistive-pulse DNA detection with a conical nanopore sensor. *Langmuir* **2006**, *22*, 10837–10843. [[CrossRef](#)]
40. Thangaraj, V.; Lepoitevin, M.; Smietana, M.; Balanzat, E.; Bechelany, M.; Janot, J.-M.; Vasseur, J.-J.; Subramanian, S.; Balme, S. Detection of short ssDNA and dsDNA by current-voltage measurements using conical nanopores coated with Al<sub>2</sub>O<sub>3</sub> by atomic layer deposition. *Microchim. Acta* **2016**, *183*, 1011–1017. [[CrossRef](#)]
41. Kececi, K.; San, N.; Kaya, D. Nanopore detection of double stranded DNA using a track-etched polycarbonate membrane. *Talanta* **2015**, *144*, 268–274. [[CrossRef](#)] [[PubMed](#)]
42. Kaya, D.; Dinler, A.; San, N.; Kececi, K. Effect of Pore Geometry on Resistive-Pulse Sensing of DNA Using Track-Etched PET Nanopore Membrane. *Electrochim. Acta* **2016**, *202*, 157–165. [[CrossRef](#)]
43. Giambianco, N.; Fichou, Y.; Janot, J.-M.; Balanzat, E.; Han, S.; Balme, S. Mechanisms of Heparin-Induced Tau Aggregation Revealed by a Single Nanopore. *ACS Sens.* **2020**, *5*, 1158–1167. [[CrossRef](#)] [[PubMed](#)]
44. Zhang, Y.; Chen, X.; Wang, C.; Roozbahani, G.M.; Chang, H.-C.; Guan, X. Chemically functionalized conical PET nanopore for protein detection at the single-molecule level. *Biosens. Bioelectron.* **2020**, *165*, 112289. [[CrossRef](#)]
45. Namuduri, S.; Narayanan, B.N.; Davuluru, V.S.P.; Burton, L.; Bhansali, S. Review—Deep Learning Methods for Sensor Based Predictive Maintenance and Future Perspectives for Electrochemical Sensors. *J. Electrochem. Soc.* **2020**, *167*, 37552. [[CrossRef](#)]
46. Gonzalez-Navarro, F.F.; Stilianova-Stoytcheva, M.; Renteria-Gutierrez, L.; Belanche-Muñoz, L.A.; Flores-Rios, B.L.; Ibarra-Esquer, J.E. Glucose Oxidase Biosensor Modeling and Predictors Optimization by Machine Learning Methods. *Sensors* **2016**, *16*, 1483. [[CrossRef](#)]
47. LaPierre, N.; Egan, R.; Wang, W.; Wang, Z. De novo Nanopore read quality improvement using deep learning. *BMC Bioinform.* **2019**, *20*, 552. [[CrossRef](#)]
48. Smith, M.A.; Ersavas, T.; Ferguson, J.M.; Liu, H.; Lucas, M.C.; Begik, O.; Bojarski, L.; Barton, K.; Novoa, E.M. Barcoding and demultiplexing Oxford Nanopore native RNA sequencing reads with deep residual learning. *BioRxiv* **2019**. [[CrossRef](#)]
49. Zeng, J.; Cai, H.; Peng, H.; Wang, H.; Zhang, Y.; Akutsu, T. Causalcall: Nanopore Basecalling Using a Temporal Convolutional Network. *Front. Genet.* **2019**, *10*, 1332. [[CrossRef](#)]
50. Landry, M.; Winters-Hilt, S. Analysis of nanopore detector measurements using Machine-Learning methods, with application to single-molecule kinetic analysis. *BMC Bioinform.* **2007**, *8*, 12. [[CrossRef](#)]
51. Winters-Hilt, S.; Landry, M.; Akeson, M.; Tanase, M.; Amin, I.; Coombs, A.; Morales, E.; Millet, J.; Baribault, C.; Sendamangalam, S. Cheminformatics methods for novel nanopore analysis of HIV DNA termini. *BMC Bioinform.* **2006**, *7*, 22. [[CrossRef](#)] [[PubMed](#)]
52. Vercoutere, W.; Winters-Hilt, S.; Olsen, H.; Deamer, D.; Haussler, D.; Akeson, M. Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat. Biotechnol.* **2001**, *19*, 248–252. [[CrossRef](#)] [[PubMed](#)]



53. Nivala, J.; Mulrone, L.; Li, G.; Schreiber, J.; Akeson, M. Discrimination among protein variants using an unfoldase-coupled nanopore. *ACS Nano* **2014**, *8*, 12365–12375. [[CrossRef](#)] [[PubMed](#)]
54. Liu, Q.; Fang, L.; Yu, G.; Wang, D.; Xiao, C.-L.; Wang, K. Detection of DNA base modifications by deep recurrent neural network on Oxford Nanopore sequencing data. *Nat. Commun.* **2019**, *10*, 2449. [[CrossRef](#)] [[PubMed](#)]
55. Wescoe, Z.L.; Schreiber, J.; Akeson, M. Nanopores discriminate among five C5-cytosine variants in DNA. *J. Am. Chem. Soc.* **2014**, *136*, 16582–16587. [[CrossRef](#)]
56. Lepoitevin, M.; Jamilloux, B.; Bechelany, M.; Balanzat, E.; Janot, J.-M.; Balme, S. Fast and reversible functionalization of a single nanopore based on layer-by-layer polyelectrolyte self-assembly for tuning current rectification and designing sensors. *RSC Adv.* **2016**, *6*, 32228–32233. [[CrossRef](#)]
57. Bayley, H.; Martin, C.R. Resistive-Pulse Sensing—From Microbes to Molecules. *Chem. Rev.* **2000**, *100*, 2575–2594. [[CrossRef](#)]
58. Luo, L.; German, S.R.; Lan, W.-J.; Holden, D.A.; Mega, T.L.; White, H.S. Resistive-pulse analysis of nanoparticles. *Annu. Rev. Anal. Chem. (Palo Alto Calif.)* **2014**, *7*, 513–535. [[CrossRef](#)]
59. DeBlois, R.W.; Bean, C.P. Counting and Sizing of Submicron Particles by the Resistive Pulse Technique. *Rev. Sci. Instrum.* **1970**, *41*, 909–916. [[CrossRef](#)]
60. Holden, D.A.; Watkins, J.J.; White, H.S. Resistive-pulse detection of multilamellar liposomes. *Langmuir* **2012**, *28*, 7572–7577. [[CrossRef](#)]
61. Yang, L.; Yamamoto, T. Quantification of Virus Particles Using Nanopore-Based Resistive-Pulse Sensing Techniques. *Front. Microbiol.* **2016**, *7*, 1500. [[CrossRef](#)] [[PubMed](#)]
62. Larkin, J.; Henley, R.Y.; Muthukumar, M.; Rosenstein, J.K.; Wanunu, M. High-bandwidth protein analysis using solid-state nanopores. *Biophys. J.* **2014**, *106*, 696–704. [[CrossRef](#)] [[PubMed](#)]
63. Storm, A.J.; Storm, C.; Chen, J.; Zandbergen, H.; Joanny, J.-F.; Dekker, C. Fast DNA translocation through a solid-state nanopore. *Nano Lett.* **2005**, *5*, 1193–1197. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).