

令和元年9月2日現在

機関番号：14401

研究種目：基盤研究(S)

研究期間：2014～2018

課題番号：26220603

研究課題名（和文）トンネル電流による1分子シーケンシング法

研究課題名（英文）Single-Molecule Sequencing Methods via Tunneling Current

研究代表者

谷口 正輝（Taniguchi, Masateru）

大阪大学・産業科学研究所・教授

研究者番号：40362628

交付決定額（研究期間全体）：（直接経費） 139,700,000円

研究成果の概要（和文）：1分子を流れるトンネル電流の計測により、DNAとRNAの塩基配列と、ペプチドの部分アミノ酸配列を決定することに成功し、1つのウイルスの全ゲノム配列決定に成功した。さらに、疾病マーカーである化学修飾塩基分子と化学修飾アミノ酸の1分子識別とともに、DNAに挿入された抗がん剤の直接観察に成功した。また、溶液に含まれる2種類のDNA・RNAの塩基配列と存在比と、2種類のペプチドのアミノ酸配列と存在比を決める定量分析法を開発した。

研究成果の学術的意義や社会的意義

1分子シーケンシング法は、現在の技術では直接解読できない、RNAの塩基配列やペプチドのアミノ酸配列を直接決定できるため、新たな生命現象の解明に寄与すると期待される。さらに、安価で高速にゲノム配列を決定できる本手法は、ゲノム解析に基づく個別化医療に寄与すると期待される。また、DNA中のがんマーカーや抗がん剤を直接観察できる本手法は、抗がん剤の作用機序を調べ、新たな抗がん剤の開発に寄与すると期待される。

研究成果の概要（英文）：In this study, we successfully concatenated the base molecules in DNA and RNA and the partial amino acid sequence of peptide and sequenced the entire genome of a virus by measuring the tunneling current flowing through a single molecule. We successfully observed the direct incorporation of our target anticancer drug into DNA and identified the individual molecules associated with the chemically modified base molecule and the amino acid, both of which are vital disease markers. A quantitative method was developed to analyze the base sequences and abundance ratios for the two types of DNA and RNA contained in the target solution. We determined the amino acid sequences and the abundance ratios for the two types of peptides under investigation.

研究分野：1分子科学

キーワード：1分子科学 量子シーケンサー DNA RNA ペプチド

1. 研究開始当初の背景

2001年に完了したヒトゲノム計画は、遺伝子に基づく個別化医療の幕開けと期待されていたが、一人の全ゲノムを解読する莫大なコストと時間が、個別化医療の大きなハードルとなった。この状況をいち早く予測した米国国立衛生研究所(NIH)は、1日と1000ドルでヒトゲノム解読を実現するDNAシーケンシング法の開発を開始した。その究極の原理は、DNAの塩基配列をトンネル電流で読取るものであり、大学・企業を含め全米体制で研究開発が行なわれた。しかし、1本鎖DNAの直径1nmに対応するナノギャップ電極を作ることの困難さのため、ナノギャップシーケンシング法の開発は不可能と考えられていた。

2010年、我々は、開発・発展させてきた1分子計測技術を用いて、1nmのナノギャップ電極を作製し、世界に先駆け、トンネル電流による1塩基識別を実現した(Nat. Nanotech. 2010.)。さらに、2012年、DNAと、既存のDNAシーケンシング法では直接読取ることが出来ないマイクロRNA(miRNA)の塩基配列決定に成功(Sci. Rep. 2012)し、ナノギャップシーケンシング法は、量子力学を使った唯一のDNA・RNAシーケンシング法として世界で認知されるに至った(図1. NewsFocus in Science 2012)。

ゲノムに基づく個別化医療を実現するためには、DNAの塩基配列解読はもちろんのこと、がん等の疾病マーカーとなるmiRNAの塩基配列解読と、ゲノムの設計図を基に作られるタンパク質のアミノ酸配列解読が必須となる。我々は、DNAとmiRNAの塩基配列決定には成功したが、20種類のアミノ酸からなり、複雑な分子構造を持つタンパク質のアミノ酸配列決定には、タンパク質を10分子程度のアミノ酸から構成されるペプチドに分割する場合でも、複雑な電流シグナルを解析する方法と、1分子構造を直線形にする1分子技術の開発を待たなければならなかった。

我々は、DNAやmiRNAを構成する全塩基の1分子電気伝導度ヒストグラムをデータベースとして、電流シグナルに各塩基分子を確率的に帰属させる情報理論(隠れマルコフモデル)を応用することで、複雑な電流シグナルを解析する方法を開発(Sci. Rep. 2012)した。さらに、ナノギャップ電極近傍の電気泳動力と電気浸透力の電圧制御により、1分子構造を直線状にする技術の開発に成功(Sci. Rep. 2011. Sci. Rep. 2012)した。また、数nmギャップの固定電極を90%以上の高い再現性で作製するプロセス(重ね電極プロセス)を確立し、1分子技術における最大の壁を突破することに成功した。これら3つの科学技術の開発により、1分子ペプチドシーケンシング法を開発する道が拓けてきた。さらに、20種類の分子を識別する1分子ペプチドシーケンシング法は、これまで実現できなかったDNA修飾・RNA修飾の1分子解像度マッピング法へと展開できると期待された。

2. 研究の目的

本研究では、ペプチドのアミノ酸配列を、トンネル電流により解読する1分子ペプチドシーケンシング法を開発するとともに、ペプチドシーケンシング法の検出原理の学理を確立する(図2)。また、タンパク質の活性化・不活性化を決定するアミノ酸のリン酸化の1分子識別と、1分子ペプチド上におけるリン酸化部位の1分子解像度の識別を行う。さらに、開発する1分子シーケンシング法を、遺伝子機能をスイッチする修飾DNAと、タンパク質合成系を正確かつ円滑に動作させる修飾RNAにおける修飾塩基分子の1分子解像度マッピング法へと発展させる。本研究の達成により、遺伝情報がDNA RNA タンパク質の順で伝達されるセントラルドグマに関わる生体分子の1分子シーケンシング法を開発する。

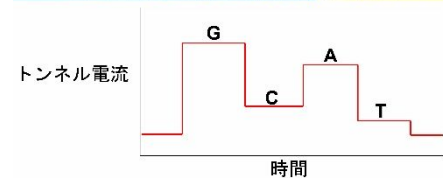
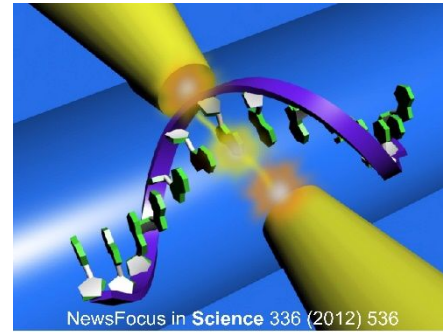


図1. トンネル電流によるDNAの塩基配列決定の原理図.

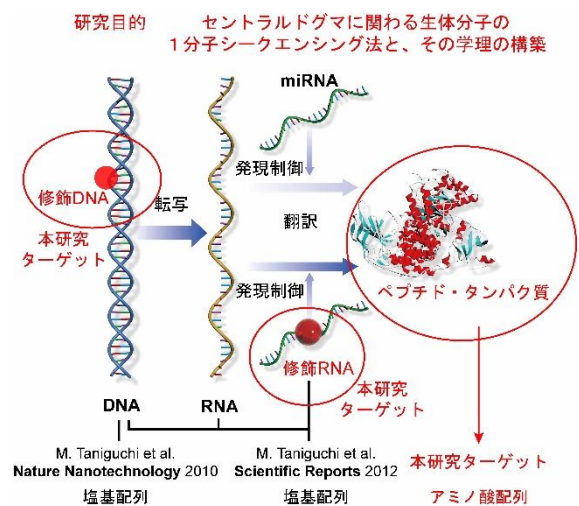


図2. 本研究の目的.

3. 研究の方法

本研究は、大阪大学産業科学研究所・谷口研究室と研究協力者から成る研究体制で実施された。得られた電流 時間波形が、ナノデバイスの構造、計測システム、および解析方法に強く依存するため、研究代表者の総括の下、谷口研究室の中で、ナノデバイス、微小電流計測、および解析の三位一体の研究を進めた。谷口研究室では、量子化学を用いた分子電子状態計算、分子を流れる電子輸送計算、流体力学、電磁気学、およびイオン輸送の連成方程式で記述されるマイクロ・ナノ流路内の1分子の流動ダイナミクスシミュレーション、および統計処理を用いて電流 時間波形解析を行った。これらの計算、シミュレーション、および解析では、実験で得られる電流 時間波形を十分に説明できない部分があるため、各分野の研究協力者とともにそれぞれの方法・理論を発展させた。特に、実験で得られる電流 時間波形に直接影響する水溶液中の1分子の流動ダイナミクスと、機械学習を用いた電流 時間波形解析を集中して研究を進めた。

4. 研究成果

本研究の目的とした、DNA・RNA上の修飾塩基分子の1分子解像度マッピングと、ペプチドのアミノ酸配列決定とリン酸化アミノ酸の1分子識別を達成した。さらに、DNAシーケンサー開発の重要なマイルストーンの一つであるバクテリオファージの全ゲノム解析に成功した。また、2種類のDNA、RNA、ペプチドの混合溶液において、DNA・RNAの塩基配列と存在比、およびペプチドのアミノ酸配列と存在比を決定する定量解析法を開発した。計測で得られるトンネル電流 時間波形の機械学習により、1つの波形データから1塩基分子を高精度で識別する手法の開発は、当初予測しなかった研究成果である。この新手法は、夾雑物中におけるターゲット分子を検出・識別できる革新的な分析手法になると期待される。

(1) 1分子シーケンサーの原理解明

1分子を介してナノ電極間に流れるトンネル電流(I)は、電圧が V のとき、次式により記述されることが理論的に示されていた。

$$\frac{I}{V} = \frac{2e^2}{h} \frac{4\Gamma_L\Gamma_R}{(E_F - E_{HOMO})^2}$$

ここで、 e と h は、電気素量とプランク定数であり、 E_F と E_{HOMO} は、電極のフェルミ準位と分子の最高占有分子軌道(HOMO)である。また、 Γ_L と Γ_R は、分子と左電極、分子と右電極との相互作用の大きさである。本研究提案当時、DNAとRNAを構成するそれぞれ4つの塩基分子の1分子電気伝導度ヒストグラムにおけるピーク伝導度が、DNAではグアニン>アデニン>シトシン>チミン、RNAではグアニン>アデニン>シトシン>ウラシルの順番であり、この順番が、各塩基分子のHOMOのエネルギー準位と同じであることを明らかにしていた。しかし、それぞれの4つの塩基分子では、分子 電極間相互作用の大きさが異なるため、1分子電気伝導度の解明には、分子 電極間相互作用が同程度であり、HOMOのエネルギーが異なる塩基分子が必要であった。さらに、分子長に対して指数関数的に減少するトンネル電流の比較は、分子長が同程度であることが前提となる。これら3つの性質を満たす分子の開発は、これまで成功例が無く、試行錯誤の結果、アデニンとグアニン上の窒素・炭素原子を炭素・窒素原子に置換した分子を設計・合成した。開発した分子の1分子電気伝導度を計測した結果、1分子電気伝導度ヒストグラムにおけるピーク伝導が、 $(E_F - E_{HOMO})^2$ に反比例することを見出した。これにより、ようやく1分子上の電子輸送は、理論提案されていた量子伝導機構であることを明らかにした(ASC Nano 2019)。

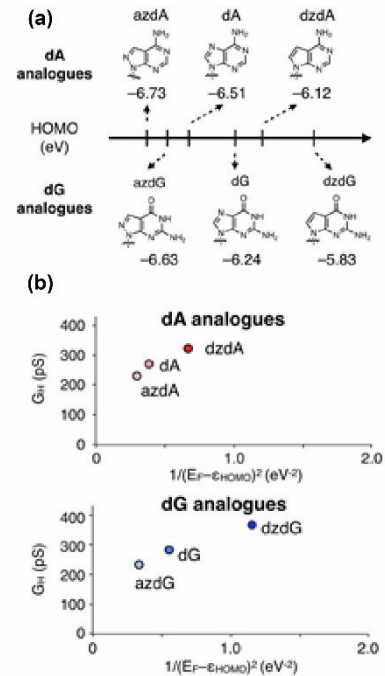


図 3. (a)開発した塩基分子類縁体と、(b)1分子電気伝導度とHOMOの関係。

(2) バクテリオファージの全ゲノム解析

全てのDNAシーケンサーの開発には、2つの共通なマイルストーンがあり、1つは5,386塩基から構成されるバクテリオファージPhiX174の全ゲノム解析であり、もう1つはヒトの全ゲノム解析である。特に、1つ目のマイルストーンは、新技術が実用化されるかを判定する重要な基準となっている。本研究では、PhiX174の全ゲノムを60塩基に断片化した部分配列を1分子シーケンサーにより解読し、塩基読取エラー率1%の部分配列をアセンブル法で結合することで全ゲノム解析に成功した(図4論文作成中)。さらに、120塩基に断片化した部分配列を用いた実験でも、全ゲノム解析に成功した。全配列解析の結果、1分子DNAにおける有意な最大読取長と最頻度読取長は、それぞれ30塩基と8塩基であり、平均読取速度は 5.24×10^6 塩基/日で

あった。また、全ての部分配列で、DNA 両末端配列の読取頻度が相対的に低いことが明らかとなり、ナノギャップ電極間への 1 分子 DNA の流動ダイナミクスが関与していることが示唆された。

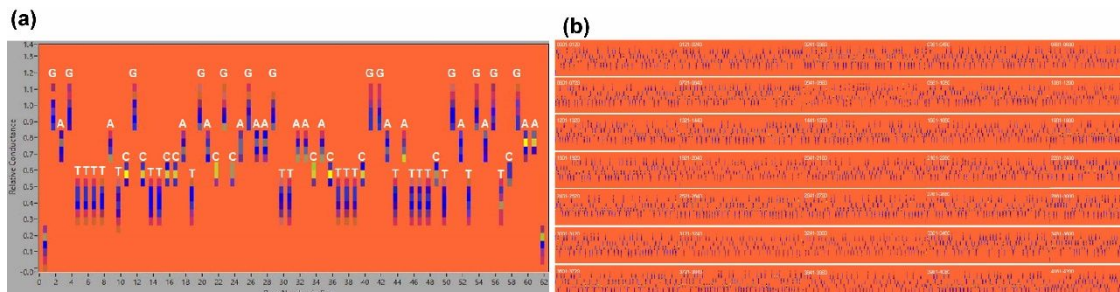


図 4. 1 分子シーケンサーにより得られた (a)60 塩基配列と (b) PhiX174 の全ゲノム配列(5,786 塩基).

(3) DNA・RNA 上の修飾塩基分子の 1 分子解像度マッピング miRNA (let7 ファミリー) の 1 分子解像度マッピングと定量解析

miRNA では、塩基配列が 1 つ異なると、異なる疾病マーカーになることが知られており、診断には 1 分子解像度マッピングが必須となる。そこで、本研究では、がんの予防・予後診断への応用が期待されている miRNA の塩基配列決定を 1 分子シーケンシング法により行った。がんマーカーとして知られる let7 ファミリーのうち、22 塩基から構成される 4 種類 (let7a、let7c、let7e、let7f) の miRNA の電流 時間波形を計測したところ、DNA と同様なスパイク状の電流シグナルが観察された。塩基分子の 1 分子電気伝導度ヒストグラムに基づき、電流 時間波形を解析したところ、多くの断片配列が得られ、7 塩基以上の連続配列を用いて miRNA の全塩基配列が決定された (図 5. Sci. Rep. 2018)。

正確ながん診断を行うためには、miRNA の塩基配列と存在比を決定できる定量解析が必要である。1 分子シーケンス法の計測では、1 つの部分配列が 1 つの DNA に対応するため、部分配列のシグナル数比 = 溶液中に存在する分子数比、の関係が成り立つと考えられる。この原理を検証するため、let7a:let7f = 3:1 で混合した溶液を用いて、1 分子シーケンスを行い、let7a と let7f の 12 番目の塩基分子のカウント数比を求めたところ、let7a:let7f = 2.6:1 が得られた。

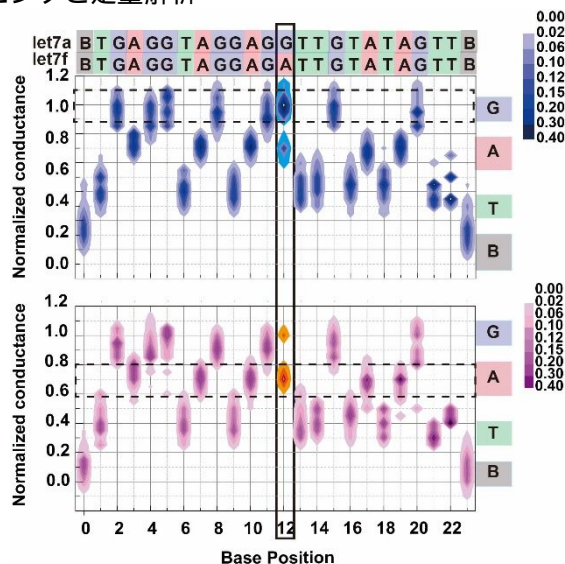


図 5. 1 分子シーケンスにより得られた let7a と let7f の塩基配列.

RNA 上の化学修飾塩基分子の 1 分子解像度マッピング

シトシンのメチル化に代表される塩基分子の化学修飾 (エピジェネティック修飾) は、遺伝子機能をオンにするスイッチの役割を持つことが明らかになっており、塩基配列に加えて、エピジェネティック修飾を同時に解読する DNA・RNA シーケンシング法が、精密医療を実現するコア技術になると考えられている。しかし、現在の DNA シーケンサーは、エピジェネティック修飾を直接解読することは原理的に困難である。

本研究では、メチル化アデニンとメチル化シトシンを導入したがんマーカーである RNA の塩基配列決定とエピジェネティック修飾の 1 分子識別を行った。メチル化アデニンとメチル化シトシンの 1 分子電気伝導度が、グアニンの 1 分子電気伝導度の 0.91 倍、1.21 倍であることを用いて、塩基配列を決定した。図 6 に示すように、メチル化アデニンとメチル化シトシンを 1 分子で識別し、全塩基配列決定に成功した (論文作成中)。

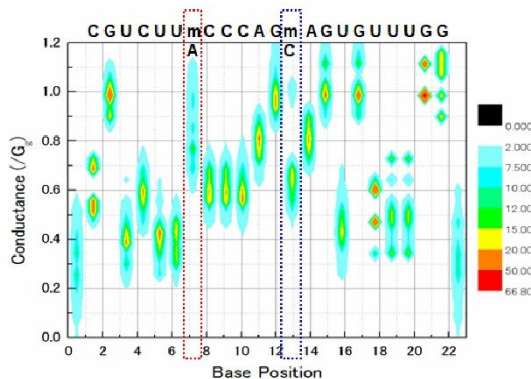


図 6. 1 分子シーケンスにより得られた化学修飾塩基を持つ RNA の塩基配列と、化学修飾塩基分子の 1 分子解像度マッピング.

メチル化アデニンとメチル化シトシンを 1 分子で識別し、全塩基配列決定に成功した (論文作成中)。

DNA 上の抗がん剤の 1 分子解像度マッピング
1 分子シーケンス法の 1 分子解像度マッピング能力を抗がん剤に適用した。抗がん剤であるフルオロリジン(FTD)は、消化器系がん用いられており、その作用機序は、複製過程においてチミンの代わりに FTD が DNA 中に挿入されることで、たんぱく質の合成が阻害され、がん細胞の増殖が抑制されると考えられている。しかし、DNA 中への FTD の挿入を直接観察した例はなく、作用機序は未だ明らかにされていない。本研究では、DNA 中の FTD の 1 分子識別を行い、抗がん剤を第 5 の塩基分子としてモデル DNA の全塩基配列決定を行った。

FTD の 1 分子電気伝導度を決定したところ、グアニンの 1 分子電気伝導度の 0.15 倍程度であった。これは、フッ素置換により HOMO のエネルギーが低下したことに対応している。得られた 1 分子電気伝導度を用いて、1 分子シーケンス法により、21 塩基からなる DNA と、塩基配列中央の 2 つのチミンを FTD で置換した DNA の塩基配列決定を行った。5 種類の分子に対応して、5 階調の 1 分子電気伝導度が観察され、塩基配列の決定とともに FTD の直接観察に世界で初めて成功した(図 7. Sci. Rep. 2019)。

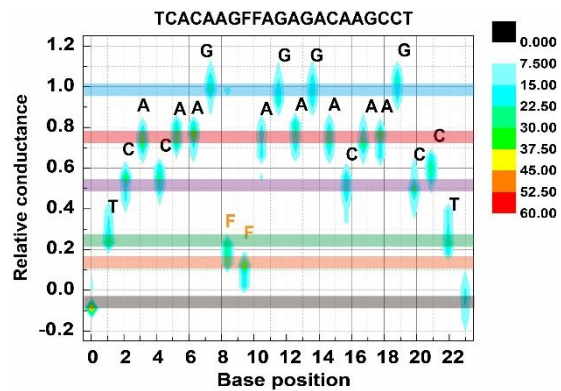


図 7. 1 分子シーケンスにより得られた抗がん剤を含む DNA の塩基配列。

(4) ペプチドの部分アミノ酸配列決定と修飾アミノ酸の 1 分子識別

アミノ酸と修飾アミノ酸であるリン酸化チロシンの 1 分子電気伝導度ヒストグラムを決定するため、0.5nm と 0.7nm のナノギャップ電極により、1 種類のアミノ酸分子のみが溶解した水溶液の電流 時間波形を計測した。いずれのナノギャップ電極でも全てのアミノ酸分子で電気シグナルが得られたが、1 分子電気伝導度ヒストグラムに明確な 1 つのピーク電流が得られたのは、0.5nm で 9 種類、0.7nm で 9 種類であり、全部で 13 種類のアミノ酸が 1 つのピーク電気伝導度を示した(Nat. Nanotechnol. 2014)。チロシンとリン酸化チロシンは、0.7nm のナノギャップ電極で識別され、チロシンの 1 分子電気伝導度が大きいことが分かった。その後、ナノギャップ電極間を安定保持することで、20 種類のアミノ酸の 1 分子識別が可能となった。

10 個のアミノ酸から構成されるペプチドとチロシンがリン酸化された修飾ペプチドの電流 時間波形を 0.7nm のナノギャップ電極で計測した。得られた電流 時間波形を、アミノ酸の 1 分子電気伝導度ヒストグラムを用いて解析したところ、アミノ酸の部分配列を決定することができ、修飾・非修飾ペプチドの識別に世界で初めて成功(図 8 Nat. Nanotechnol. 2014)した。

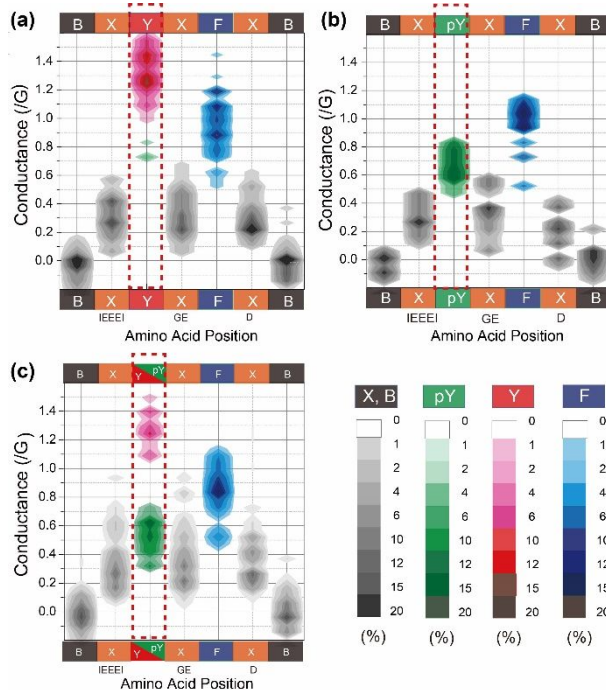


図 8. 1 分子シーケンスにより得られたペプチドの部分アミノ酸配列と、リン酸化チロシンの 1 分子解像度マッピング。

(5) 機械学習による高精度 1 分子識別

1 分子シーケンスにおける解析法の最大の課題は、統計的に得られる 1 分子電気伝導度ヒストグラムの重なりによる読取精度の低下であった。また、トンネル電流の高感度さのため避けることが困難なナノギャップ電極間距離の変動や不純物に由来する電気ノイズを除去し、ターゲット分子から得られる 1 つの電流 時間波形のみを抽出・解析する新手法の開発も課題であった。これらの課題は、電流による 1 分子計測に共通であり、1 分子計測にも関わらず、1 分子から得られる 1 つの電流シグナルを解析出来てないという本質的な課題であった。

高い再現性・安定性を持つ 1 分子シーケンスシステムの開発と、これにより得られる高い再現性の大量データ、さらに 1 分子電気伝導機構の解明により、1 つの電流 時間波形が、分子の HOMO、分子構造、および分子 電極間相互作用の分子固有の情報を持つことが分かった。これら

の情報を波形データから引き出すことができれば、1つの波形データから1分子を高精度に識別できると考え、波形データから1分子固有の情報を引き出す機械学習を用いた解析法を開発した。特に、避けることが困難なノイズ除去法の開発では、個々のノイズを学習し、原因解明するのではなく、複数のノイズを1群のノイズとして学習することで、ノイズ+ターゲット分子シグナルのデータから、ノイズを除去する(Positive Unlabeled Classification)PUC法を用いた(図9 J. Phys. Chem. C. 2019)。

開発した解析法を用いると、1つの電流時間波形で、4種類の塩基分子をF値=0.55~0.91の高い精度で識別することに成功した。さらに、3'-AT-5'と5'-AT-3'も高い精度で識別することができる事を見出し、DNA・RNAにおける3'端と5'端を識別できる驚くべき結果が得られた。このように、機械学習を用いた解析法の開発は、1分子計測の課題を解決するとともに、予測しなかった結果をもたらした。

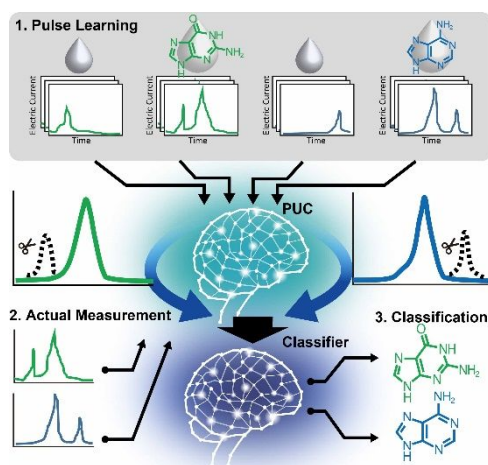


図9. PUC法による高精度1分子識別法。

5. 主な発表論文等

〔雑誌論文〕(計27件)

Takafumi Furuhashi, Takahito Ohshiro, Gaku Akimoto, Ryosuke Ueki, Masateru Taniguchi, and Shinsuke Sando, Highly Conductive Nucleotide Analogue Facilitates Base-Calling in Quantum-Tunneling-Based DNA Sequencing, ACS Nano, 査読有, Vol. 13, No. 5, 2019, pp. 5028-5035, DOI: 10.1021/acsnano.9b01250.

Takahito Ohshiro, Makusu Tsutsui, Kazumichi Yokota, Masateru Taniguchi, Quantitative Analysis of DNA with Single-Molecule Sequencing, Scientific Reports, 査読有, Vol.8, 2018, pp.8517(1-8), DOI: 10.1038/s41598-018-26875-7.

Massimiliano Di Ventra and Masateru Taniguchi, Decoding DNA, RNA and Peptides with Quantum Tunneling, Nature Nanotechnology, 査読有, Vol. 11, No. 2, 2016, pp. 117-126, DOI: 10.1038/nnano.2015.320.

Takahito Ohshiro, Makusu Tsutsui, Kazumichi Yokota, Masayuki Furuhashi, Masateru Taniguchi, Tomoji Kawai, Detection of Post-Translational Modifications in Single Peptides Using Electron Tunneling Currents, Nature Nanotechnology, 査読有, Vol.9, No.10, 2014, pp.835-840, DOI: 10.1038/nnano.2014.193.

〔学会発表〕(計64件)

〔図書〕(計4件)

〔産業財産権〕

出願状況(計5件)

〔その他〕

ホームページ等

<http://www.bionano.sanken.osaka-u.ac.jp/>

6. 研究組織

(1)研究協力者

研究協力者氏名: 筒井 真楠

ローマ字氏名: TSUTSUI Makusu

研究協力者氏名: 大城 敬人

ローマ字氏名: OHSHIRO Takahito

研究協力者氏名: 横田 一道

ローマ字氏名: YOKOTA Kazumichi

研究協力者氏名: 小本 祐貴

ローマ字氏名: KOMOTO Yuki