

JST news

未来をひらく科学技術

特集

データの共有と統合で 生命科学に変革を

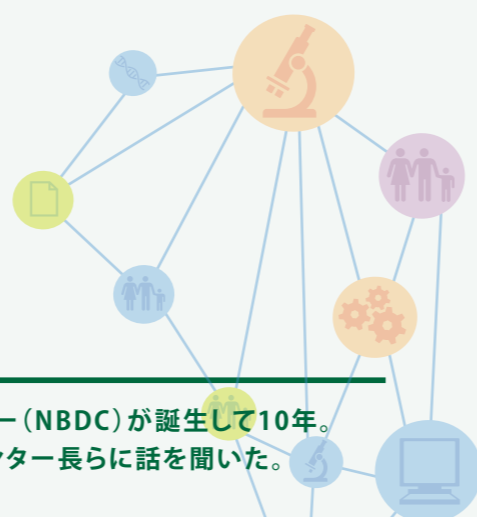
12

December
2020



特集

データの共有と統合で 生命科学に変革を



生命科学分野のデータベース統合を目指すバイオサイエンスデータベースセンター (NBDC) が誕生して10年。これまでの成果や研究を取り巻く環境の変化、今後の展望について、高木利久センター長らに話を聞いた。



データが価値を生む ビッグデータ時代の到来

生命科学では50年に一度の周期で大きな変革が起きている、とNBDC前運営委員長である堀田凱樹博士は言う。その言葉通り、1900年にはメンデルの法則の再発見、53年にはDNA二重らせん構造の発見が生命科学に大きなパラダイムシフトをもたらした。そして2003年のヒトゲノムの解読は生命科学の研究スタイルを大きく変え、情報生物学の時代が幕を開けた。これを先取りする形で、01年JSTはバイオインフォマティクス推進センター (BIRD) を立ち上げ、生物学の情報技術開発、データベース構築、人材育成を目指した。

その後の10年で、ゲノム解読装置の性能が1万倍になるなど計測技術は革命的といえるほど進歩し、データ量は指数関数的に増加した。人工知能やスーパーコンピューターも驚くほど進歩した。これにより、さまざまな種類の

データの組み合わせから人工知能を使って新たな仮説を生み出すという研究アプローチが生まれた。これを「データ駆動型科学」と呼ぶ。データそのものが価値を生むビッグデータ時代が訪れたのである。

研究者は自分の実験で得たデータだけでなく、他の研究者のデータもより積極的に必要とし始めた。しかし当時は多くのデータが研究室や研究プロジェクトの中に囲い込まれ、他人のデータを活用するのは容易でなかった。

データを世の中に解き放ち、誰でも自由に使えるようにすることが必要だった。複数のデータベースの組み合わせから仮説を得るには、公開するだけでなく各データベースのフォーマットや専門用語を統一しなければならない。これがデータベース統合プロジェクトの始まりである。このプロジェクトは06年頃から文部科学省を中心に生まれ、その推進センターとして11年にNBDCが設立された。

データを使いやすく提示 研究者の利便性向上

「世界には2万以上の生命科学データベースがあるとされますが、フォーマット、専門用語、精度などがばらばらで実験研究者はもとより、バイオインフォマティクスの専門家でも多様で膨大なデータを使いこなすには高いスキルが必要です。NBDCではデータの共有と統合を通じて研究者の負担を大幅に減らし、データベースの利便性を高める努力をしてきました」とNBDCの高木利久センター長は言う。

例として高木センター長は13年に開始したNBDCヒトデータベースの構築を挙げる。これは疾患研究のために主に日本人のゲノムデータなどを集めたものである。データの登録数は年々増え続け、疾患研究に欠かせないデータベースになった(図1)。

NBDCで企画運営に携わる眞後俊幸主査は、データ統合の具体例を挙げる。これまでに報告された遺伝子のオンオフに関わる文献に用いられた膨大なデータは、公共のデータベースに登録され世界中の誰でも見ることができるという。そこでNBDCの支援で京都大学の沖真弥特定准教授はそれらのデータを整理統合し、見やすく表示した「ChIP-Atlas」を開発したところ、公開からわずか5年で150以上の研究論文で引用されたという。「データがどこかにあって使えるというだけでは駄目なのです。それらを整理統合して初めて価値が出ます」と眞後主査は成果を強調する。

これと並行して、世界中の主要なデータベースの「RDF化」を進めてきた。

RDFはデータとその意味を主語、述語、目的語の3つ組で統一的に表現するもので、これにより多数のデータベースが網の目のようにつながり、2次的、3次のデータベースを素早く柔軟に作ったり、さまざまな視点で考察したり、人工知能を含む多様なツールを適用できるため、容易に仮説を生成できる(図2)。

例えば、「TogoVar」というデータベースはRDF化されたデータベースを用いたことにより非常に効率良く作ることができた。開発コストが低減した結果、実験研究者が求めてきたサービスを短期間で提供できるようになったのだ。

RDFデータベースは現時点ではまだ検索用語が実験研究者向けではないので、文献データベースのように研究者が日常的に使うまでには至っていない。高木センター長は「はからずも新型コロナウイルスの影響でオンライン会議が一気に普及したように、資材不足などで実験がしにくくなっている今、データベースの活用が研究推進力として重要度を増してくれば、RDFデータベースを縦横無尽に使ってデータ駆動型科学を実践する段階が訪れると思います。それまでは地道にデータベース基盤を整備していきます」と言う。

伝道と人材育成が鍵 データ駆動型科学の実践

「統合データベースができ、格段に使いやすくなったとはいえ、研究者がストレスなく使えるまでにはまだまだです。これからはデータベース作りだけでなくデー

タ駆動型科学を実践し、RDFデータベースも含めデータベースがこんなに研究に役立つんだよということを示していきたい」と高木センター長。

NBDCの守備範囲は医学、創薬、農学、有用物質生産など幅広いが、「まずは医学の分野で研究者や臨床医が日常的に使えるものを目指します」と続ける。TogoVarだけでなく、入力した塩基配列に対してゲノム編集で重要となるCRISPR-Cas9システムのガイドRNAを設計できる「CRISPRdirect」や、患者の症状を入力すると関連する希少・難治性疾患の候補を可能性が高い順に自動的にリストアップする医療者向け検索システム「PubCaseFinder」などのサービスを通じて、その目標は確実に達成されつつある。

誰でも簡単に使えるデータベースを作るには高度な専門知識が求められる。「NBDCの主な役割ではありませんがデータベース作りの面白さ、重要性を広

く伝え、この分野への人材の参入、育成にも貢献したい」と話す高木センター長自身も、データ駆動型科学の伝道者だ。

新技術との融合 30年後の生命科学とは

50年周期なら、次の変革は2050年だ。その時には何が起きるのか。「30年先のことは予想もつきませんが、10年後くらいには実験ロボットを使って網羅的で均質なビッグデータが簡単に得られ、それを人工知能で統合解析し、自動的に仮説を出すという時代になっているのではないのでしょうか」と高木センター長。

ロボットと人工知能がさらに活躍する未来に、生命科学を取り巻く状況はどうなっているのか興味は尽きない。データベース開発を通じて、生命科学研究に新潮流を生み出し続けるというNBDCの夢はこれからも続いていく。

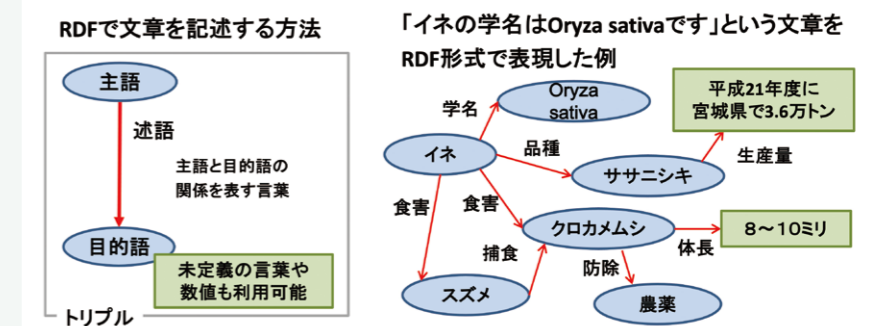


図2 RDF (Resource Description Framework) は、コンピューターが処理できるよう、さまざまな事象の「主語」と「目的語」、それらの関係を表す「述語」でつないだ「3つ組(トリプル)」表記する文法形式だ(左)。例えば「イネの学名はOryza sativaです」という文章をRDF形式で記述すると、さまざまな概念をつなげることができる(右)。RDF化によって、他分野のデータベースとの統合解析の他、コンピューターによる推論も可能になり、新しい発見につながる。

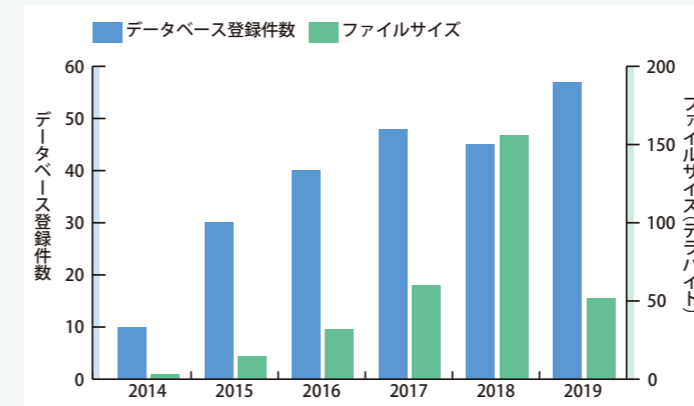


図1 NBDCヒトデータベースに登録されるヒトゲノムデータの件数は、2013年10月に運用を開始して以来大幅に増加している。



研究成果

未来社会創造事業 探索加速型
研究領域「共通基盤」
研究課題「創薬を加速する細胞モデリング基盤の構築」

誰でも使える創薬支援解析ツールを開発 1人1人に最適な治療法の選択が可能に

多様ながん治療薬が世界中で研究開発されています。がん医療は、がんの大きさや広がり具合、遺伝子発現の有無などいくつかの指標による分類と治療が主流です。しかし従来の方法では患者の状態に合わせた治療が難しく、患者によって予後に大きな違いが生じてしまうため、患者1人1人の遺伝子情報に基づいた新たな分類法が求められています。

大阪大学蛋白質研究所の岡田眞里子教授は、がんゲノムデータを用いて、薬剤応答による分類や、がんタイプに合わせて最適な治療法を選択する「個別化医療」を目的とした解析ツールを構築しました。

公共の遺伝子情報データベースから取得した網羅的ながん細胞の遺伝子発現データと、細胞中の遺伝子間の相互作用や活性を数理モデルで記述しコンピューター上で再現する細胞シミュレーションを組み合わせました。そして、がん細胞増殖の目印であり薬剤探索の指標となる、リン酸化酵素(キナーゼ)の活性を定量的に予測する計算手法を開発しました(図1)。

人工知能の一種である遺伝的アルゴリズムを使い、複数の細胞のキナーゼ活性の実験データを数理モデルに学習させ、パラメーター最適化計算を並行して行い、多数のパラメーターセットを取得します。これを別の細胞に適用すれば、その都度実験しなくても遺伝子発現データだけを用いて、細胞シミュレーションにより薬の効果を予測できます。

このような解析には熟練者でも約2~3ヵ月かかっていた。しかし

し本手法ではプログラミングや数値計算の知識や経験がなくても初心者約1週間で使えるようになります(図2)。現在、国内外のアカデミックや製薬企業など9つのプロジェクトでこの手法を用いた解析を開始しており、特許や論文などの成果が出始めています。「細胞シミュレーションと臨床公共データを組み合わせれば、これまで現場で

しか行えなかった創薬研究のデジタル化が可能になります。公共データを用いることで、初期の研究費用は数十分の1に、計算を用いることで研究開発時間は半分以上削減することができます。患者の遺伝子情報を基にシミュレーションを

行えば、予後の良し悪しを予測し、さらにその原因を探し出して薬のデータベースから最適な薬を選ぶことができます」と、岡田教授は創薬研究の進展や個別化医療の確立に向けて力強く語ります。他分野のシミュレーション解析にも応用できるため、工学など多様な分野への展開も期待されます。

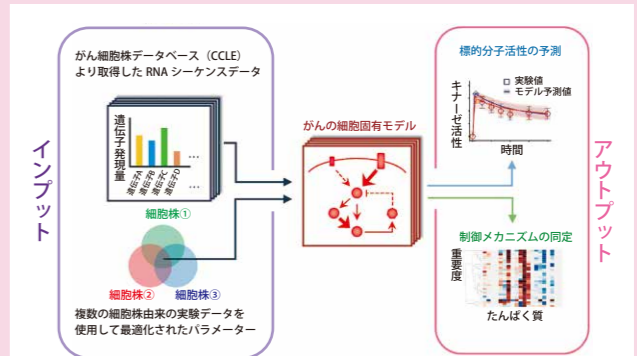


図1 細胞のリン酸化実験データを基に学習させた数理モデルに、公共データベースから取得した網羅的ながん細胞の遺伝子発現データを入力(インプット)することにより、創薬標的であるキナーゼの実活性値を予測し、その活性制御に関わる重要遺伝子を解析し同定(アウトプット)する。

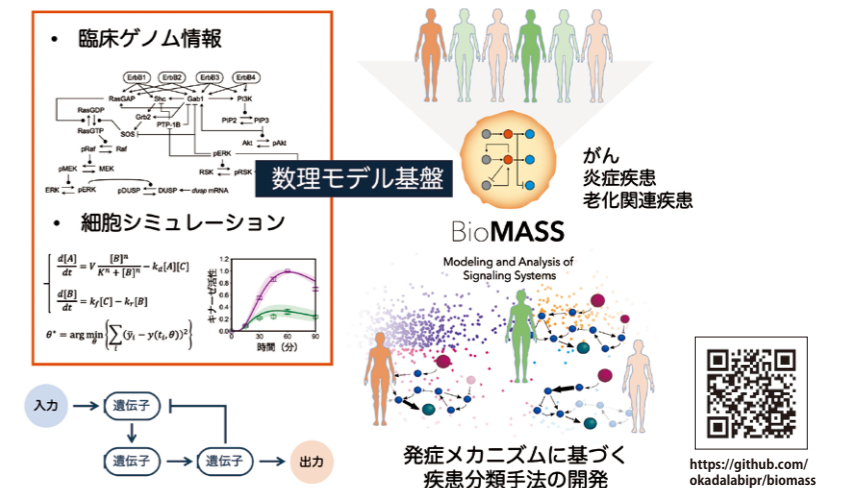


図2 開発したBioMASS(Modeling and Analysis of Signaling Systems) はオープンソースソフトウェアとして公開されている。臨床ゲノム情報と合わせることで、発症の仕組みに基づく疾患の分類や薬剤の選択が可能になる。



とよおか り ひと
豊岡 理人
NBDC 研究員

日本人ゲノム多様性 統合データベース 「TogoVar」 疾患の解明に期待

日本人のバリエーション情報を公開 論文情報もワンストップで提供

ヒトゲノム配列の解読は1990年代より欧米を中心に進められ、2003年には標準的なゲノム配列「参照配列」として公開された。同じヒトという種でも、ゲノム配列は人種や個人ごとに異なる。

NBDCは、日本人のゲノム配列の多様性に関する情報を集めたデータベースとしてTogoVarを18年6月に公開した。「参照配列との違いを表すバリエーションの出現頻度を日本人集団と他集団で比較するだけでなく、生物学的意味や関連する論文もワンストップで検索できます」とTogoVarの開発に携わる豊岡理人研究員は説明する(図3)。

希少疾患の原因を絞り込み 国内外がデータに注目

20年7月には、全ゲノム解析した日本人集団のサンプル数としては最大規模となる、7609人分のゲノム配列データを統合・再解析して得られたバリエーション情報の総合的なリストを日本の複数機関が連携して作成し、TogoVarから公開した。このデータは国内外の大学や研究所、民間企業など、公開からわずか2カ月で100近い研究グループに利用されているようだ。

子供だけに発症した希少疾患がある場合、その両親にはなく子供だけにあるバリエーションは疾患と関係する可能性がある。しかし、親子のゲノムを比較しただけでは子供に固有のバリエ

ーションの数はまだ多く、どれが疾患の原因であるかの特定は困難だ。「日本人集団の中に存在するバリエーションも子供のバリエーションから取り除ければ、疾患の原因となるバリエーションをより早く特定できます」と豊岡研究員(図4)。これまで未知・未解明だった日本人特異的な希少疾患や単一遺伝子疾患の病因解明に役立つ基礎的なデータとして期待される。

「今回のデータでは日本人固有のバリエーションがたくさん見つかっています。日本人に特異的なデータを蓄積、公開することは日本の研究者だけでなくアジアを中心に海外の研究者へも貢献できると考えています」とも豊岡研究員は話す。実際、TogoVarは50カ国からのアクセスがあり、利用されている。

研究者が使いたくなる ユーザー目線の開発

豊岡研究員はシステムエンジニアなどを経て、遺伝子解析の道に進み、TogoVarの開発に関わるようになった。

「タイで結核の疾患関連遺伝子研究のためにゲノム解析をしていたことがありますが、今取り組んでいる仕事は、当時の自分が欲しかったデータベースを作り出すことです。TogoVarの開発では、使う側の立場に立って、インターフェースも使いやすく工夫しました」と振り返る。

今後はゲノムのバリエーション情報だけでなく、遺伝子や疾患をまとめた情報や遺伝子発現に関する情報も追加し、さらなる利便性の向上を目指すという。

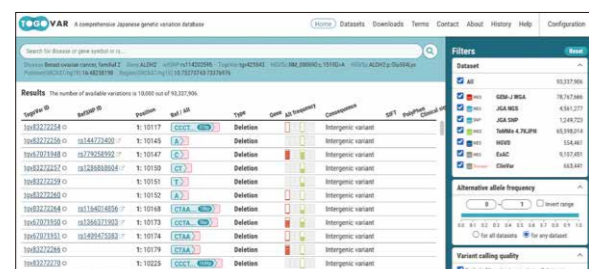


図3 TogoVarは、トップページから個々のバリエーション情報、データベースごとの頻度、疾患との関与などをワンストップで検索できる。

Step 1. 子固有のバリエーション情報を、父母子の全ゲノム配列を取得したのちに抽出
Step 2. 子固有のバリエーション情報から日本人一般に存在するバリエーションを取り除き、疾患に関連するバリエーションを特定



図4 TogoVarで提供する日本人のバリエーション情報を使い、日本人特異的な希少疾患や単一遺伝子疾患の病因解明を促進する。



戦略的創造研究推進事業CREST
 研究領域「素材・デバイス・システム融合による革新的ナノエレクトロニクスの創成」
 研究課題「超高速・超低電力・超大面積エレクトロクロミズム」

カーテンやブラインドに代わる調光ガラス 耐久性や操作性の実証実験がスタート



図1 取り付けが済んだ調光ガラス(左)と設置した窓の外観(右)。

まぶしい直射日光を避けたり、寝る時に外の光を防いだりするために、私たちはカーテンやブラインドを使用します。しかし、部屋が広く窓が多いビルや施設ではカーテンやブラインドの開け閉めに手間がかかります。

この手間を省き、遮光状態と透明状態をスイッチ1つで自由に切り替えられるのがエレクトロクロミック(EC)調光ガラスです。EC調光ガラスは航空機のボーイング787の窓などで使用されているものの、消費電力や発色性に課題があり、普及が進んでいませんでした。

物質・材料研究機構機能性材料研究拠点の樋口昌芳グループリーダーは、青、緑、紫、黒など多彩な色を出して発色性も良く、低消費電力で駆動する新しいEC材料「メタロ超分子ポリマー」の開発と実用化に向けた研究を進めています。今年6月には東京化成工業(東京都中央区)と共同で、メタロ超分子ポリマーの安定供給を可能にする合成プロセスを確立し、一般販売を開始しました。これらの成果を受け9月8日からスタートしたのが、調光

ガラスとして建物に設置した状態での耐久性や操作性などの実証実験です。ベンチャー企業育成や市内研究機関の技術の活用を推進する茨城県つくば市の支援によって実現したもので、来年3月までの予定でつくば市のつくばスタートアップパークで行われています。

実験では、青色と紫色のメタロ超分子ポリマーで作られた10センチメートル四方のEC調光ガラスを1000枚以上用意し、木枠に入れて既存の窓の内側に取り付けました(図1、図2)。「公共施設なので、誰でも来て操作することができます。ぜひ多くの方にお越しいただいて未来のガラスを体感してもらいたいです」と樋口グループ

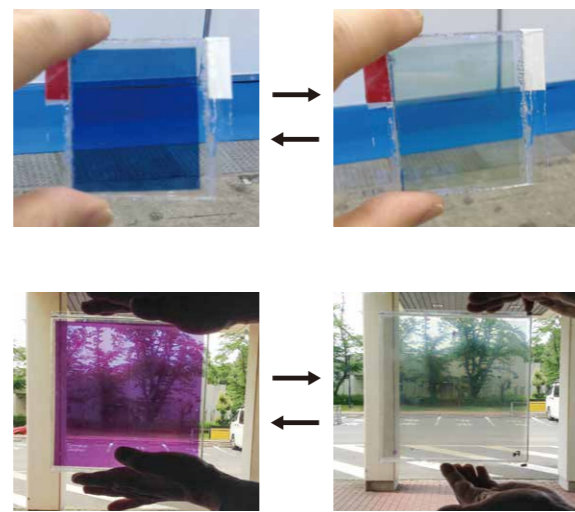


図2 青色(上)と紫色(下)のメタロ超分子ポリマーで作製したEC調光ガラスにおける色変化の様子。電圧をかけると発色と透明が切り替わり、切り替えた後は電源を切ってもその状態が維持される。

リーダーはコメントしています。

同様の実証実験は神奈川県川崎市でも行うことが決まり、来年2月から設置準備が始まる予定です。多くの声を参考に、実用化に向けた改良を進めるとともに、これらの実証実験を通じた製造業者への技術紹介と供給網の構築も目指し取り組んでいます。



戦略的創造研究推進事業CREST
 研究領域「実験と理論・計算・データ科学を融合した材料開発の革新」
 研究課題「実験・計算・データ科学融合による塗布型電子材料の開発」

有機半導体をシャボン膜方式で塗布 薄膜トランジスタの特性が理論限界に迫る

印刷技術によって軽量で柔らかな電子回路やデバイスを構築する「プリントエレクトロニクス」。常温常圧の条件下、インクを塗って溶剤を乾かすだけでデバイスを構築できるため、有機半導体による薄膜トランジスタの実現も期待されています。しかし、性能向上に適したフッ素樹脂などの絶縁層の表面上ではインクが強くはじかれ丸くなってしまい、半導体結晶膜の形成に必要な薄く均質な液膜の形成は困難でした。

東京大学大学院工学系研究科の長谷川達生教授らは、シャボンが枠に薄い液膜を張る仕組みにヒントを得て、インクを強くはじく基板上でも溶液先端の湾曲(メニスカス)を延ばして均質な半導体結晶膜を形成できる



図1 シャボン膜は枠に付着して薄い液膜を保持している。

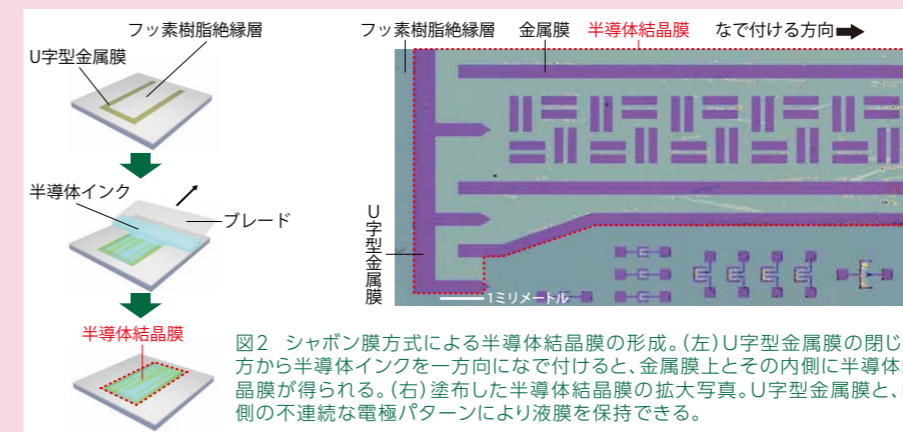


図2 シャボン膜方式による半導体結晶膜の形成。(左)U字型金属膜の閉じた方から半導体インクを一方になで付けると、金属膜上とその内側に半導体結晶膜が得られる。(右)塗布した半導体結晶膜の拡大写真。U字型金属膜と、内側の不連続な電極パターンにより液膜を保持できる。

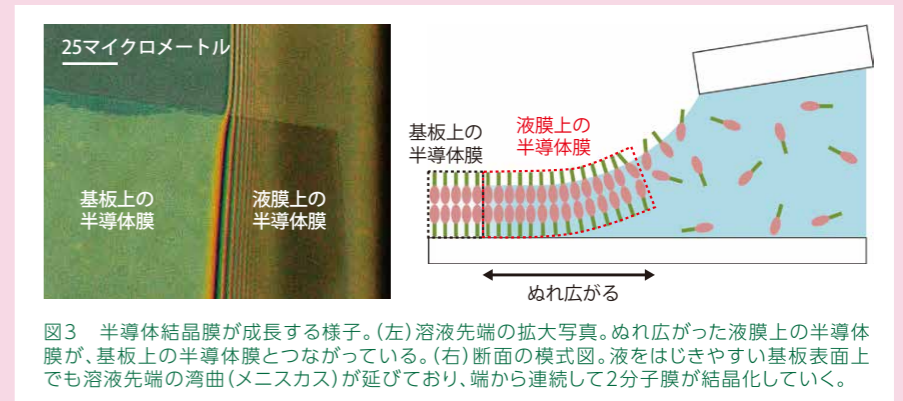


図3 半導体結晶膜が成長する様子。(左)溶液先端の拡大写真。ぬれ広がった液膜上の半導体膜が、基板上の半導体膜とつながっている。(右)断面の模式図。液をはじきやすい基板表面上でも溶液先端の湾曲(メニスカス)が延びており、端から連続して2分子膜が結晶化していく。

「拡張メニスカス塗布法」を開発しました(図1)。鍵を握る半導体材料は棒状の有機分子、フェニルベンゾチエノナフトチオフェンです。フッ素樹脂による絶縁層上にU字型の金属膜の枠を形成し、半導体インクをU字の閉じた方から一方向になで付けると、薄さわずか2分子分の半導体膜が金属膜上とその内側にも形成されました(図2)。

膜形成の様子を顕微鏡で観察したところ、U字型の枠で三方を囲われた領域内では溶液がはじかれずにぬれ広がっていました(図3)。シャボン膜の端が枠に付着し、その内側で薄い液膜が保持される現象と似ています。

できた半導体膜を調べた結果、有機

分子が秩序良く並んだ均質な単結晶で構成されていることが確認されました。また、この成膜法で薄膜トランジスタを作製し電気特性を評価したところ、2ボルト以下の低電圧で安定して駆動し、室温動作での理論限界に近いオンオフ応答を示しました(図4)。全て常温常圧の塗布工程により、優れた薄膜トランジスタの開発に成功しました。

今後は実用化に向けてさらなる高性能化を目指し、材料の探索と塗布法の高度化を進めていきます。電子デバイスの軽量化やウェアラブル化、そして省資源化を実現する次世代半導体として期待されます。

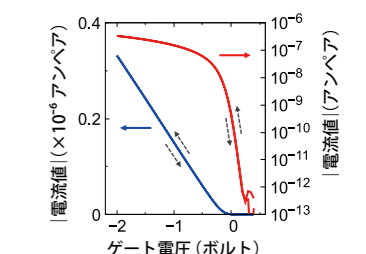
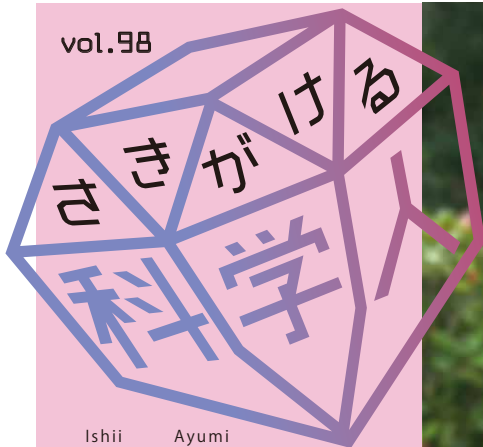


図4 作製した有機薄膜トランジスタの伝達特性。2ボルトで駆動、ドレイン電圧を0.2ボルトに固定して測定した。線形(左軸)を対数(右軸)で表したときの勾配が急峻で、オンオフ応答が良好なことを示している。

vol.98



Ishii Ayumi
石井 あゆみ

JST さきがけ研究者
桐蔭横浜大学 大学院工学研究科
特任講師



見えない光を捉えたい

Profile

神奈川県出身。2008年 青山学院大学大学院理工学研究科博士課程修了。博士(理学)。ソニー先端マテリアル研究所研究員、桐蔭横浜大学大学院工学研究科助教、青山学院大学理工学部助教を経て、18年より現職。17年よりさきがけ研究者。

Q1. 研究テーマを選んだきっかけは？

A1. 無機と有機が両方好きで

横浜大好きハマっ子です。小さい頃は工作をしたり、鉱物や化石を集めたりするのが好きでした。高校の実験で還元された銀が析出する反応に感動し、大学では化学を専攻することに。有機と無機が両方好きで、錯体化学はそれらのハイブリッドです。「全く異なる性質を組み合わせれば、素晴らしいものになるはず」。長谷川美貴教授の研究室の1期生となって希土類錯体の研究を始め、この分野では誰にも負けまいと必死で勉強しました。

現在は、有機と無機を界面で化学的に融合したハイブリッド材料の光機能を生かして、微弱な光や近赤外光、



偏光といった「目に見えない光」を高感度に捉えるフォトダイオードを開発しています。偏光は電場や磁場の振動方向がそろった光で、有機分子が持つキラリティによって無機物質にらせん性を持たせた構造を利用すると直接検出できます。今まで見えなかった光を捉えるセンサーを実現し、さまざまな現象の解明に貢献したいです。

Q2. 研究者としての心構えは？

A2. 出合いを大切に

異なるものが混ざるとすごいのは、さきがけの「光極限」研究領域も同じです。ただくっつけるのではない、本当の異分野融合ですね。良い研究仲間ができて、コロナ禍でもオンラインでつながっています。

「明るく楽しく元気良く」をモットーに、出合いを大切にしてきました。桐蔭横浜大学の宮坂教授のお陰で今の進展がありますし、かつて研究生としてお世話になった北海道大学の山崎巖教授の研究室では、山崎トモ子先生との出会いが衝撃でした。圧倒的な洞察力とスピードに「こうなりたい」と強く思いました。研究者への志を固くしたのはその時です。

Q3. 出産・育児の乗り切り方は？

A3. さきがけと保育園と周りの理解

任期雇用の研究者は研究と並行して次の職探しをします。最も手を動かしたい時期で、育休は取りにくい。私はさきがけに採択された時、2人目の妊娠がわかりました。さきがけにはライフイベントを支援する制度があり、中断した分を上限に研究期間を延長できます。産休、育休を取ることができたので、研究と子育てが両立しました。若手も安心して子供を産める社会になってほしいですし、そのためには雇用環境の改善が急務だと思います。

コロナ禍でも、保育園がしっかりと予防対策をしてくれるので仕事に集中できます。お迎えや子供が熱を出した時には何がなんでも帰らなければなりません、理解してくれる周りの支えに感謝しています。

休日は近所の公園でとにかく芝生を走り回ったり、滑り台で遊んだり、ドラえものの映画を見たりしています。娘たちも、石ころやどんぐりをポケットに詰めていますね。



JSTは、シンクタンク機能、研究開発、産学連携、次世代人材育成、科学と社会との対話など、多岐にわたる事業を通じて、持続可能な開発目標(SDGs)の達成に積極的に貢献していきます。



編集 長：安孫子満広
科学技術振興機構(JST)広報課
制 作：株式会社伝創社
印刷・製本：株式会社丸井工文社

リサイクル適性(A)

この印刷物は、印刷用の紙へリサイクルできます。



JSTnews

December 2020

発行日/令和2年12月4日
編集発行/国立研究開発法人科学技術振興機構(JST)総務部広報課
〒102-8666 東京都千代田区四番町5-3サイエンスプラザ
電話/03-5214-8404 FAX/03-5214-8432
E-mail/jstnews@jst.go.jp JSTnews/https://www.jst.go.jp/pr/jst-news/



最新号・バックナンバー