

戦略的プロジェクト研究推進事業

「民間事業者等の種苗開発を支える『スマート育種システム』の開発」

令和元年度 最終年度報告書

中課題番号	19314849
中課題名	育種を加速するパスイ型シミュレータの開発とバイオデータ連携基盤構築

研究実施期間	令和元年度～令和2年3月31日
代表機関	国立研究開発法人 理化学研究所
研究開発責任者	和田 智之
研究開発責任者 連絡先	TEL : 048-467-9827
	FAX :
	E-mail : swada@riken.jp
共同研究機関	株式会社電通
	学校法人慶應義塾 慶應義塾大学SFC研究所
	SyntheticGestalt株式会社
	学校法人早稲田大学 ナノ・ライフ創新研究機構
	bitBiome株式会社
普及・実用化 支援組織	

<別紙様式3>最終年度報告書

I-1. 年次計画

研究課題	研究年度		担当研究機関・研究室	
	令和元年度		機関	研究室
1 育種ターゲットの把握の加速化に係る研究開発 (1) ロボットを活用したバイオデータ生産拠点 (AOI-PARCバイオラボ(パイロット版)) の構築  (2) AIを活用した演繹的アプローチに基づく、腸内マイクロバイオームのバイオメカニズムモデルの構築  (3) (2) で構築したモデルに基づくマイクロバイオームの健全化に資する有用成分の探索  (4) 有用成分生産に適した作物の育種を可能とするパスウェイ型シミュレータの開発  (5) (4) で開発したシミュレーションおよび生産されたバイオデータのバイオデータ連携基盤への接続	バイオラボの構築	←→	電通京都ビジネスア クセラレーションセ ンター 理研	事業共創部  光量子制御技術開発チ ーム
	有用成分の探索	←→	理研	光量子制御技術開発チ ーム
	パスウェイ型シミュ レータの開発	←→	Synthetic Gestalt	
	APIの開発	←→	Synthetic Gestalt	
	WAGRIの機能拡張 データ取引市場構築	←→	慶應義塾大学 SFC 研究所	政策・メディア研究科
2 既存データ有効活用に係る研究開発 (6) 最新のクラウドやネットワーク関連技術等を活用し、多様なバイオ関連研究データを相互に連携、分析を可能とするバイオデータ連携基盤の構築  (7) バイオ分野におけるデータ利用規約(案)の策定  (8) 非活用データの連携・活用に向けたユースケース開発	専門家ヒアリング ELSE文献調査		慶應義塾大学 SFC 研究所	政策・メディア研究科
	ユースケース開発		慶應義塾大学 SFC 研究所	政策・メディア研究科
3 腸内マイクロバイオームを対象としたシングルセルゲノム解析システムの開発 (9) シングルセルゲノム解析システムの構築  (11) 腸内細菌の時系列解析  (12) ゲノムデータの情報連携基盤との共有	大規模なゲノム情報 解析場の構築		早稲田大学  bitBiome	竹山研究室
	ゲノムデータ時系列 解析		早稲田大学 bitBiome	竹山研究室
	APIの開発	←→	早稲田大学 bitBiome	竹山研究室

## I-2. 実施体制

研究項目	担当研究機関・研究室		研究担当者
	機関	研究室	
研究開発責任者	理化学研究所	光量子光学研究センター 光量子制御技術開発チーム	◎和田 智之
1. 育種ターゲットの把握に関する研究開発	電通	京都ビジネスアクセラレーションセンター事業共創部	○志村 彰洋
(1) ロボットを活用したバイオデータ生産拠点 (AOI-PARC バイオラボ(パイロット版))の構築	電通	京都ビジネスアクセラレーションセンター事業共創部	○志村 彰洋
(2) AIを活用した演繹的アプローチに基づく、腸内マイクロバイオームのバイオメカニズムモデルの構築	理化学研究所	光量子光学研究センター 光量子制御技術開発チーム	△和田 智之
(3) (2)で構築したモデルに基づくマイクロバイオームの健全化に資する有用成分の探索	理化学研究所	光量子光学研究センター 光量子制御技術開発チーム	△和田 智之
(4) 有用成分生産に適した作物の育種を可能とするパスイエイ型シミュレータの開発	Synthetic Gestalt	研究開発部	島田 幸輝
(5) (4)で開発したシミュレーションおよび生産されたバイオデータのバイオデータ連携基盤への接続	Synthetic Gestalt	研究開発部	島田 幸輝
2. 既存データ有効活用に係る研究開発	慶應義塾大学SFC研究所	環境情報学部	○神成 淳司
(6) 最新のクラウドやネットワーク関連技術等を活用し、多様なバイオ関連研究データを相互に連携、分析を可能とするバイオデータ連携基盤の構築	慶應義塾大学SFC研究所	政策・メディア研究科	△信朝 裕行
(7) バイオ分野におけるデータ利用規約(案)の策定	慶應義塾大学SFC研究所	政策・メディア研究科	△信朝 裕行
(8) 非活用データの連携・活用に向けたユースケース開発	慶應義塾大学SFC研究所	政策・メディア研究科	△信朝 裕行
3. 腸内マイクロバイオームを対象としたシングルセルゲノム解析システムの開発	早稲田大学	竹山研究室	○ 竹山春子
(9) シングルセルゲノム解析システムの構築 (竹山春子、細川正人、西川洋平)	早稲田大学	竹山研究室	○ 竹山春子 △ 細川正人 △ 西川洋平
(10) 腸内細菌の時系列解析	bitBiome		△ 細川正人 △ 有川浩司
(11) ゲノムデータの情報連携基盤との共有	早稲田大学 bitBiome	竹山研究室	○ 竹山春子 △ 細川正人

(注1) 研究開発責任者には◎、小課題責任者には○、実行課題責任者には△を付す。

中課題番号	19314849	中課題 研究期間	令和1年度
大課題名	民間事業者等の種苗開発を支える「スマート育種システム」 の開発		
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデ ータ連携基盤構築		
代表機関・研究開発責任者名	国立研究開発法人理化学研究所 光量子工学研究センター 光量子制御技術開発チーム チームリーダー 和田 智之		

### I-3. 研究目的

農業競争力強化プログラムでは、今後、良質かつ低廉な種苗を開発・供給し、我が国の農業競争力の強化を図ることとされている。一方、バイオテクノロジーの発展により、農作物のゲノム情報等の育種に関するデータが蓄積され、これを用いた新たな育種技術が登場している。また、育種（品種開発）を加速するためには、「育種技術の高度化」だけでなく、どのような品種を開発するか（育種ターゲット）を判断するための技術の開発も必要となる。

このため、本研究では、

1. 育種ターゲットの把握の加速化に係る研究開発
2. 既存データ有効活用に係る研究開発
3. 腸内マイクロバイオーームを対象としたシングルセルゲノム解析システムの開発

により、

1. ヒトの腸内マイクロバイオーームに着目した少なくとも1つの特定症状または傾向の改善に寄与する成分の特定につながるバイオメカニズムモデルの構築
2. 植物の代謝メカニズムのシミュレータを構築し、バイオメカニズムモデルで特定した成分を効率的に生成するための育種ターゲットの探索
3. （非活用データを含めた）多様なバイオ関連研究データを連携可能なバイオデータ連携基盤の構築
4. 腸内マイクロバイオーームの時系列ゲノムデータの効率的な取得を可能にする解析システムを構築

を目標とする。

その結果、

1. 国内の多数の研究機関が参加することで、多様な未活用データの活用が図られることによる研究開発の加速化
2. バイオ由来機能性食品や食品会社等の民間事業者からの研究開発投資が促進され、結果としての国民の健康増進への大きな貢献

が期待される。

## I-4. 研究方法

本事業では、育種（商品開発）の加速を実現するために、（1）育種ターゲットの把握に関する調査研究、（2）既存データの有効活用に係る研究開発、（3）腸内マイクロバイームを対象としたシングルセルゲノム解析システムの開発という3つの小課題を設定して、研究を推進した。

### （1）育種ターゲットの把握に関する研究開発

「アウトカム発想に基づく育種」を実現するため、健康への関心の高まりから今後重要性を増してくるヒトの腸内マイクロバイームに着目し、マイクロバイームの健全化に資する成分を特定する手法を開発するとともに、植物の代謝メカニズムのシミュレーションを通じて当該有効成分を効率的に作る育種ターゲットを特定するシステムを開発した。

そのため、「(1) ロボットを活用したバイオデータ生産拠点 (AOI-PARCバイオラボ(パイロット版)) の構築」では、必要な実験・計測環境の要件を確認するため、モデルとなる生産制御過程を想定して、実際に実験プロトコルを走らせた。具体的には、シミュレータで単糖類の生産制御過程（単糖の生産比率にフォーカス）に関する仮説を作り、バイオラボ（パイロット版）で実際に遺伝的・環境的条件にあった育種、培養を行うことで、仮説検証を行うことができるサイクルを構築した。

また、「(2) AIを活用した演繹的アプローチに基づく、腸内マイクロバイームのバイオメカニズムモデルの構築」では、腸内細菌に関する約18,000件のサンプルから抽出した遺伝子情報を解析した細菌叢情報、ならびにサンプル提供者の個人特性、食生活、健康データを統合しバイオデータ連携基盤上に登録する方法を検討し、APIを構築することで、AIなど各種分析に利用できるようにした。また、それら情報に基づき実施される多面的な相関的解析結果から得られる知見を体系的に統合化し、特定の目的の下で活用できる仕組みを検討、バイオデータ連携基盤のAPIを構築し、一部の知見を抽出、登録することで、バイオメカニズムモデルを構築した。

「(3) (2) で構築したモデルに基づくマイクロバイームの健全化に資する有用成分の探索」では、このバイオデータ連携基盤に構築したバイオメカニズムモデルAPIを利用し、健康に関連する要件を抽出し、それらの要件に寄与する有効成分の探索を行った。

さらに、「(4) 有用成分生産に適した作物の育種を可能とするパスウェイ型シミュレータの開発」では、ある成分を生産するための植物選択と育種技術を開発するために、パスウェイ（植物内の物質の生産過程）を微分方程式でモデル化し、遺伝子因子・環境因子の両面から改善する方針を立て、育種にフィードバックをかけるシミュレータを開発した。ある関心事に対して定義される目的関数について（例えば累積収量など）、シミュレーションを通じて得られる予測値を元に、パスウェイに関して遺伝的または環境的な介入仮説を決定できるようにした。ウェット実験における遺伝子のノックアウトやノックダウン・栄養源投入などを、シミュレーションでは微分方程式を構成する変数や係数の追加や削除、値の固定や動的変更などで表現することによって、Digital Twin として機能するシミュレータとして開発した。

「(5) (4) で開発したシミュレーションおよび生産されたバイオデータのバイオデータ連携基盤への接続」では、開発したシミュレータをデータ連携基盤のもとで展開し、研究

者が利用できるようにするため、シミュレータを他のバイオインフォマティクスのツールと同様に、データに有用な情報を付加する装置として位置づけ、プラグイン形式にて、バイオデータ連携基盤の一部として取り込ませることによって実装するとともに、シミュレータと基盤互換なスキーマを定義した。

## (2) 既存データの有効活用に係る研究開発

生物学的現象等のバイオデータの収集・活用を促進するため、現状散在している未活用データを含めた多くのバイオ関連データを、柔軟に連携して運用することを可能とするバイオデータ連携基盤を構築した。

そのため、「(6)最新のクラウドやネットワーク関連技術等を活用し、多様なバイオ関連研究データを相互に連携、分析を可能とするバイオデータ連携基盤の構築」では、SIP第I期「次世代農林水産業創造技術」において、慶應義塾が中心となり構築した「農業データ連携基盤(WAGRI、以下WAGRI)」の成果を活用し、バイオデータ連携基盤に必要とされる新たな機能(バイオデータ連携基盤向け機能、セキュリティ機能、データ取引市場)を研究開発することによって、これらを統合した「バイオデータ連携基盤」を構築した。

「(7)バイオ分野におけるデータ利用規約(案)の策定」では、WAGRIの運用規則として本研究グループが中心となり農林水産省とも連携して作成済みのデータ利用規約を踏まえて、新たなバイオ分野におけるデータ利用規約(案)を、各省から発表済みのガイドラインも参照し、農林水産省並びに関係者の意見を踏まえて策定した。ELSIの観点からデータ連携基盤が具備すべき機能についても、文献調査によって最先端の情報を収集した。

「(8)非活用データの連携・活用に向けたユースケース開発」では、SyntheticGestalt社が開発するパスウェイ型シミュレータの結果と、バイオラボから生産されるデータとの連携を図ることによって、2箇所以上の研究拠点(AOI機構、理化学研究所を想定)における具体的な研究開発での利用シーンに応じた利活用への課題や研究推進における利便性等を検討し、ユースケースとして取りまとめた。なお、実際の利用は理研と利用者との契約を前提として検討した。

## (3) 腸内マイクロバイオーームを対象としたシングルセルゲノム解析システムの開発

研究の加速化に向けて、これまでに開発したシングルセルゲノム解析手法(SAG-gel法)のハイスループット化を行い、腸内マイクロバイオーームの時系列ゲノムデータの取得を可能にするシステムを構築した。

そのため、「(9)シングルセルゲノム解析システムの構築」では、ドロップレットの高速作成装置や微量分注機などの機器を導入することにより、ドロップレットの作成速度の向上およびDNA増幅後のドロップレットを用いた次世代シーケンサー用ライブラリー調製工程のハイスループット化を行い、一般実験室で再現可能なプロトコルの確立を試みた。

「(10)腸内細菌の時系列解析」では、構築したシステムのもとで、理化学研究所辨野先生のグループより提供された経時取得された糞便サンプル(連続5日間(計6サンプル))を用いて腸内マイクロバイオーームのシングルセルゲノム解析を実施した。

「(11)ゲノムデータの情報連携基盤との共有」では、Miseqから出力された個々の配列情報を元に、*de novo assembling*および生成されたContigの評価を行う解析パイプラインの構築を行うとともに、取得された大規模な配列情報を情報連携基盤に共有した。

それぞれの小課題に係る研究は、それぞれグループ単位で実施したが、定例会議の場で成果を共有し、意見交換を行うことによって、コンソーシアム全体としての成果のとりまとめにも注力した。

## I-5. 研究結果

上記の研究方法のもと、本研究を通じてウェット、ドライの研究基盤を構築し、データの取得解析を行うとともに、こうしたデータを共有できる情報連携基盤を構築することにより、育種の加速に資する以下の研究結果を得た。

### (1) 育種ターゲットの把握に関する研究開発

「(1) ロボットを活用したバイオデータ生産拠点(AOI-PARCバイオリボ(パイロット版))の構築」では、シミュレータで単糖類の生産制御過程(単糖の生産比率にフォーカス)に関する仮説を設定することを想定し、実験を試行的に実施した。先行研究においても経路が自明な多糖は少ないが、単一の単糖から構成される多糖は候補が多い。そのためシミュレーション上、多糖を構成する単糖の構成比率を変えることで、合成過程をシミュレートし、高度な研究に資する良好な結果が得られた。

また、シミュレータを用いて設定する仮説等の検証を実施可能な環境を構築した。従来よりも高度な解析が可能だけでなく、一部ロボットによる自動化により作業の効率化が実現できる環境とした。また、仮説検証を円滑に行うため、各機器をより最適に配置した。

一般的な実験/培養・測定の流れは、サンプル調整→分注→セルソート→培養→シーケンシング及び/または光学的測定、となる。この流れを円滑にすることにより、バイオリボ(パイロット版)での作業がより効率化し、仮説検証のためのバイオデータの生産性が向上することが期待できる。従来のバイオデータ生産拠点では、各機器が単体で置かれていたり、別々の場所に置かれていたり、作業の流れを阻害する物品が配置されていたりするため、本研究では、上記の一般的な実験/培養・測定の流れに沿ったレイアウトとした。



バイオリボ(パイロット版)の概観

また、出力データのデータ構造を確認し、実験結果(各機器の出力データ)に付随するメタ情報、アノテーションも確認・整理したうえで、モデル植物からの細胞培養、それら

を対象とした多種多様な成分分析、シーケンシング、イメージングの結果を一貫して管理することができるバイオデータ連携基盤への接続APIを構築した。

「(2) AIを活用した演繹的アプローチに基づく、腸内マイクロバイオームのバイオメカニズムモデルの構築」では、腸内細菌に関する約18,000件のサンプルから抽出した遺伝子情報を解析した細菌叢情報、ならびにサンプル提供者の個人特性、食生活、健康データをAIなど各種分析に利用できる形式でバイオデータ連携基盤上に登録するAPIを構築し、検証した。また、サンプルデータに対する各種相関的解析結果から得られる知見を体系的に整理統合化し、特定の目的に対し活用できる論理ネットワークとして機能する仕組みを検討、バイオデータ連携基盤のAPIを構築した。

「(3) (2) で構築したモデルに基づくマイクロバイオームの健全化に資する有用成分の探索」では、バイオデータ連携基盤に構築したバイオメカニズムモデルAPIを利用し、健康に関連する要件として以下を抽出したうえで、これらの健康要素に対し、バイオメカニズムモデルAPIを利用し、寄与する腸内細菌、栄養成分を探索し、以下の結果を得た。結果として、健康に寄与する栄養素として「食物繊維」「酪酸」「プロバイオティクス（乳酸菌、ビフィズス菌、納豆菌、発酵乳）」が抽出された。また、肥満予防への寄与する細菌種として、「バクテロイデス類」を得ることができた。

腸内細菌種・栄養成分	
酪酸菌	増加
食物繊維	摂取
フィーカリバクテリウム・プラウスニツイ（酪酸産生菌）	多い
バクテロイデス・ブルガタス	多い
コプロコッカス（酪酸産生菌）	多い
バクテロイデーテス類（バクテロイデス門）	多い
プロバイオティクス	摂取

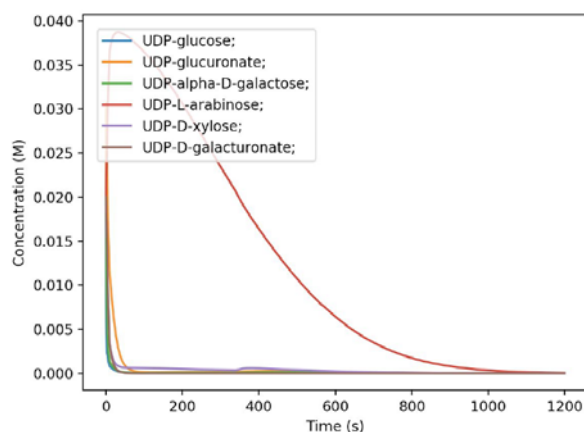
さらに、「(4)有用成分生産に適した作物の育種を可能とするパスウェイ型シミュレータの開発」では、シミュレータの構築に必要なデータを各種データベースから入手するとともに、レタスとソルガムそれぞれの実験データを用いて、データ同化を用いてパラメタの推定を行うことで土台のパスウェイを精緻化した。

具体的には、Simulator によるバーチャル実験を行うことによって、有望な介入オプション（介入仮説）を導出することができた。例えば、富 ATP 環境下や富 UTP 環境下等におけるシミュレーションを実施し、特定物質をノックアウトすることによって単糖合成比の顕著な変化が予測されることを明らかにした。さらにシミュレータの学習を通じて、シミュレータの精度を高めた。

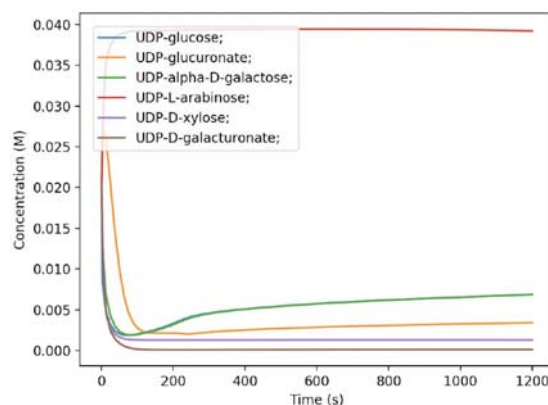
また、パスウェイからシミュレータの自動構築アルゴリズムを用いて、学習済みのシミュレータを計算機に実行可能な形式に変換し、後述するようにユーザーが利用可能なよう



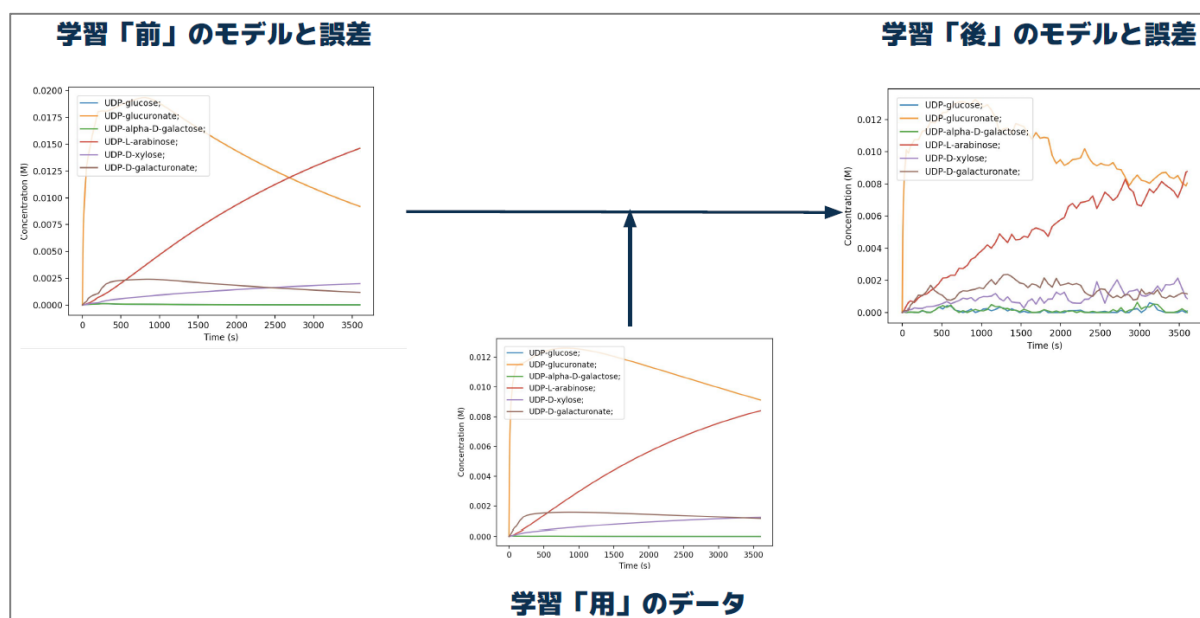
にデータ連携基盤上に展開した。



富 ATP 環境下のシミュレーション結果

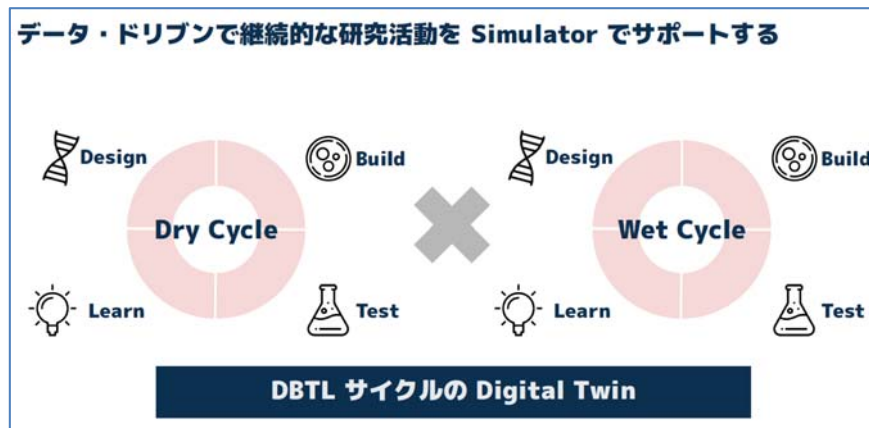


富 UTP 環境下のシミュレーション結果



学習によるシミュレータの精緻化

「(5) (4) で開発したシミュレーションおよび生産されたバイオデータのバイオデータ連携基盤への接続」では、シミュレータのバイオデータ連携基盤への接続を通じて、ユーザーが実験時と同様の入力データを用いて予測による実験結果を得られる環境を構築した。学習済みシミュレータを実行ファイル化し、それをコンテナ内に格納した上で、Web/API サーバを自動ビルドする方法を用いることによって、パスウェイで定義される全ての化合物濃度の時系列情報が出力として返却され、シームレスな形でドライとウェットの往来が可能な研究環境をバイオデータ連携基盤上に構築した。



バイオデータ連携基盤場におけるシミュレータの実装イメージ

## (2) 既存データの有効活用に係る研究開発

「(6)最新のクラウドやネットワーク関連技術等を活用し、多様なバイオ関連研究データを相互に連携、分析を可能とするバイオデータ連携基盤の構築」では、WAGRIをベースとしてバイオデータ連携基盤で重要となる以下の機能を開発した。

(1) 研究者が簡単に、効率的に利用するために必要な機能の拡張

- ・バイオ関連データに関する多様なデータフォーマット（CSV、JSON等）への対応
- ・利用者が直接的にデータを取り扱うことができるインターフェースの開発（ドラッグ&ドロップによるデータ登録機能）

(2) 研究者が安心して使えるための機能開発

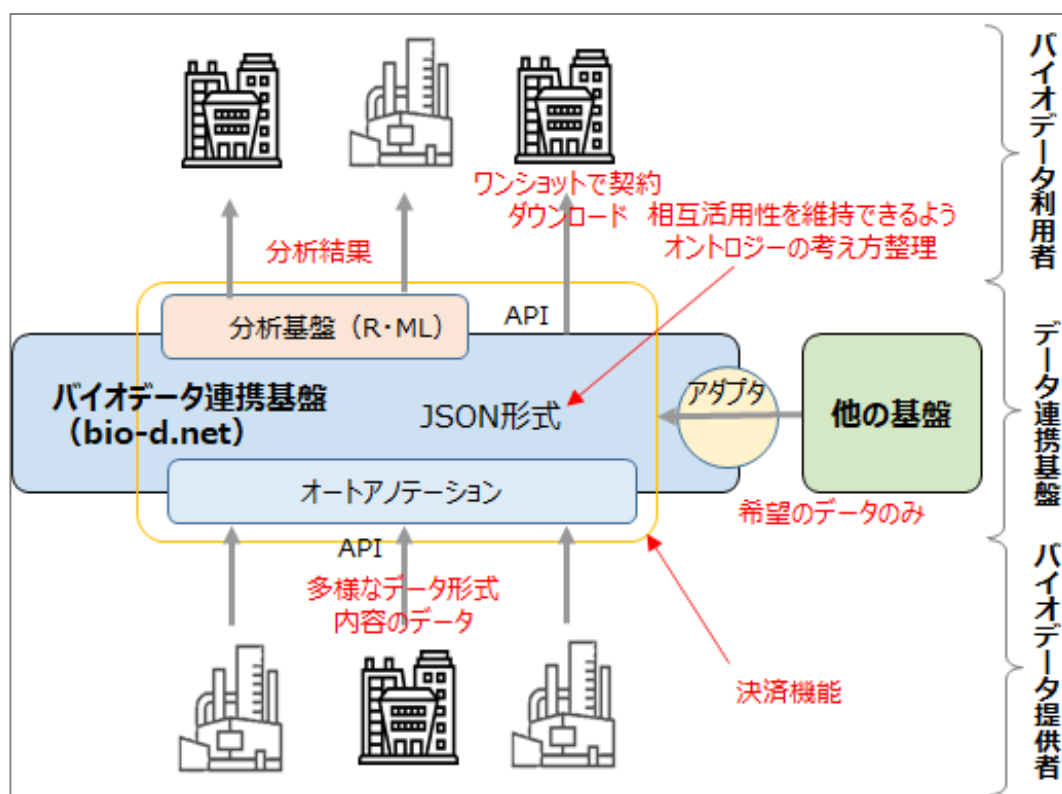
- ・WAGRIはBtoBtoCのビジネスモデルであり、法人（B）を利用者として想定しているため、データのアクセスコントロールも法人（企業ID）単位が前提となっていたが、バイオデータ連携基盤では、法人の中に所属している一研究者単位でのアクセスコントロールが必要になることも想定されるため、個人（個人ID）単位でアクセスコントロールできる機能を開発した。

(3) 研究データが積極的に活用されるために、以下の特徴を有するデータ取引市場機能を開発した

- ・会員種別（一般会員、プレミアム会員）により、データ活用範囲を定義できる機能
- ・利用者が希望するデータの検索性向上のためのマッチング機能（アノテーション、検索等）
- ・その他、データ取引に必要な決済、請求機能など

また、「(7) バイオ分野におけるデータ利用規約（案）の策定」では、これまでの各省ガイドラインを踏まえ、バイオ領域ならではの特質を加味し、さらに今後想定されるAI活用の観点からにも十分したバイオデータ利用規約（案）を策定した。また、当規約においては、データの直接的な取引のみならず、基盤を通じた分析等の利用についても対応することでデータ提供元である研究機関や研究員の懸念を低減し、実践的なデータ連携を加速しうるものとした。

「(8) 非活用データの連携・活用に向けたユースケース開発」では、18,000件の腸内マイクロバイオームデータの入手に興味のある麒麟社をユーザーとし、(6)で構築したデータ連携基盤およびデータ取引市場にて、データを利活用するユースケースを開発した。また、麒麟社へのヒアリングの結果から、対象データそのものの閲覧だけでなく、基盤上で研究者が対象データを容易に解析できる機能があると便利であるという意見を踏まえて、解析環境も導入した。



バイオデータ連携基盤の概要

### (3) 腸内マイクロバイオームを対象としたシングルセルゲノム解析システムの開発

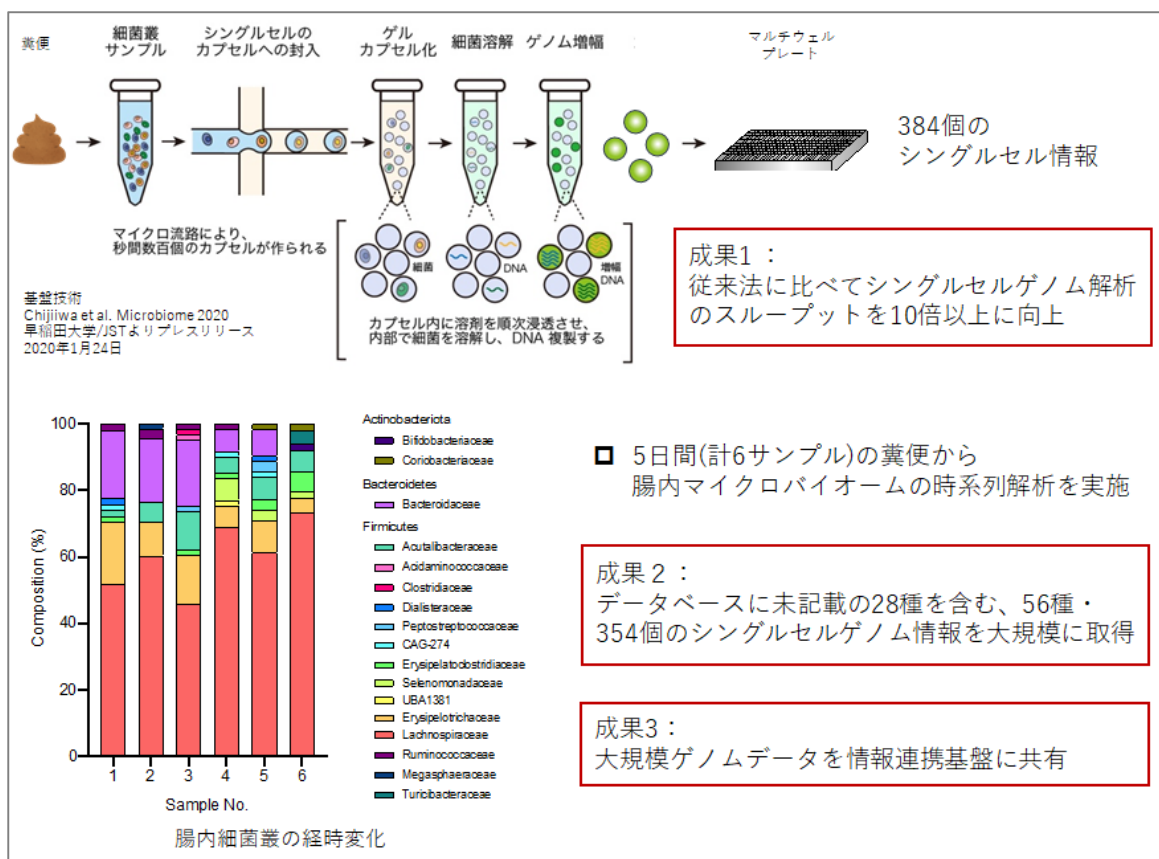
「(9) シングルセルゲノム解析システムの構築」では、機器の導入により、一度に384個のシングルセルを対象とした配列解析を可能とするシステムを構築した。本システムの構築により、一日最大768個のシングルセルゲノム解析の並行処理が可能となり、従来までの手法に比べてスループットが10倍以上に向上した。また、作業工程の自動化を実施することにより、次世代シーケンサーから出力されるサンプル毎のデータ量のばらつきが抑制さ

れ、より均質な配列情報の取得が可能になった。

「(10)腸内細菌の時系列解析」では、経時的に採取した6点分の糞便サンプルから計576サンプルのシーケンスを行った。ゲノムの品質を評価するツールであるCheckMを用いて、得られたゲノム情報の完全性(Completeness)および冗長率(Contamination)を評価した。この結果、56種354個の細菌からなるシングルセル由来のゲノム情報 (SAG: Single amplified genome) を大規模に取得できた。得られたゲノム情報は、主要な腸内細菌として知られるバクテロイデス門、ファーミキューテス門の細菌の情報を含んでおり、取得日によらず安定した菌叢を呈することが明らかとなった。

また、NCBIに登録された既知細菌のゲノムデータベースを参照した結果、28種の細菌がデータベースに記載のない新種に相当することが判明し、本研究において開発した解析システムを腸内マイクロバイオームの解析に適用することにより、未記載の細菌を含む大量のシングルセルゲノム情報が取得できることが実証された。

加えて、「(11)ゲノムデータの情報連携基盤との共有」では、得られた配列情報を情報連携基盤に共有した。また、得られたゲノムの平均長、GC%、データベース上での最近縁種とその相同性などの情報をまとめたファイルを出力し、必要に応じて参照できるシステムを構築した。



### シングルセルゲノム解析システムに係る成果概要

## I-6. 今後の課題

民間の育種を加速させるために、本研究で構築したバイオラボ（パイロット版）やシングルセルゲノム解析システムを始めとするウェット型の研究基盤や、バイオデータ連携基盤及び、そこに実装されたパスウェイ型シミュレータ等のドライ型の研究基盤を活用して出来ることを、様々な手段を用いて明示的に民間企業に周知させる必要がある。

具体的な研究開発での利用シーンに応じた課題や、研究推進における利便性等を本研究の成果として公開し、初めてバイオデータ連携基盤やバイオラボに関わる人でも、しっかりと当該拠点で出来ること/出来ないことを理解できる状況を作り出す必要がある。

また、本事業終了後は、シミュレータとバイオラボ（パイロット版）を組み合わせ、理研とAOIにて経過措置として運営するが、同時に、データ駆動型バイオ推進コンソーシアムへの民間企業参画を加速させ、当該コンソーシアムからの委託に基づき運営費用をまかない、且つ、特定の営利企業等の影響を必要以上に受けることなく、中立的な立場での継続的な運営を行うことが必要である。

参画する民間企業が増えれば、ユースケースやデータ自体が増えることに繋がり、シミュレータ自体も高度化していくので、様々な分野においてメインとなるステークホルダーに参画を依頼し、リードステークホルダーの参画を受けて業界全体の参画加速を行っていく必要がある。

加えて、様々な育種に係る実験やシミュレーションに対応するため、今回、研究基盤として構築したバイオラボ（パイロット版）及びシングルセルゲノム解析システムを参考とすることで、別の研究機関・研究場所でも複数のデータ生産拠点が構築されることが望ましい。バイオデータ連携基盤が、これら複数のバイオラボと柔軟に接続することにより、幅広い研究への対応と育種ターゲットの加速を図る必要がある。

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデータ連携基盤構築		
小課題名（実行課題名）	1 育種ターゲットの把握に関する研究開発 （1）ロボットを活用したバイオデータ生産拠点 （AOI-PARCバイオラボ(パイロット版)）の構築		
小課題 代表研究機関・研究室・研究者名	株式会社電通 京都ビジネスアクセラレーションセンター 事業創造部 志村彰洋		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

育種ターゲットを推定するため、植物内物質変化を追跡するパスウェイ型シミュレータにおける、各反応係数を予測するための実験結果を提供するデータ連携拠点として、ロボットを活用したバイオデータ生産拠点(AOI-PARCバイオラボ)のパイロット版を構築する。

育種に対する介入オプションとしては、バイオラボ(パイロット版)における実験では、遺伝子のノックアウト/ノックダウン、環境因子としての栄養源投入など、シミュレーション上では、微分方程式を構成する変数や係数の追加・削除、値の固定や動的変更などで表現されるため(Digital Twinとして機能)、バイオラボ(パイロット版)上では、柔軟な実験環境と計測環境を構築する必要がある。

実験・計測環境の構築においては、既存システムと新たなに構築する環境を効率的に組み合わせながら、下記に示す、4つの実験的アプローチを可能とする環境を構築する。

- ① 遺伝子の改変、環境因子の制御に対する植物のレスポンスの短時間での計測系を、異なる環境条件下で複数を並列して育種試験を実施する。
- ② 超臨界クロマトグラフィーや、GCMASS、LCMASSを対象に、植物に生産される物質の計測をロボット化し、複数のサンプルを並列して計測を実施する。
- ③ 植物におけるマテリアル生成過程を数理モデルと比較するため、ゲノムからの単細胞におけるマテリアル生成過程等、シンプルな条件下での合成生物学的な実験を行う。
- ④ マイクロバイームが生産する、健康に資する物質の生産過程やこれらの物質を生み出す細菌叢を育成する成分の特定

加えて、構築するバイオラボ(パイロット版)環境をバイオデータ連携基盤に接続し、モデルとなる実験プロトコルを実施することで、今後新たに発生してくると想定される、

バイオデータ連携基盤やシミュレータ型からのリクエストを、実験・計測環境に随時反映していくために必要な各種データスキーマや要件を確認する。

## 2) 研究方法

必要な実験・計測環境の要件を確認するため、モデルとなる生産制御過程を想定して、実際に実験プロトコルを走らせる。具体的には、シミュレータで単糖類の生産制御過程（単糖の生産比率にフォーカス）に関する仮説を作り、バイオラボ（パイロット版）で実際に遺伝的・環境的条件にあった育種、培養を行うことで、仮説検証を行うことができるサイクルを構築する。

バイオラボ（パイロット版）では、モデル植物もしくはモデル植物から抽出して培養した細胞に対して、大別して、成分分析、シーケンシング、イメージングが可能な環境を整えることで、研究目的の4つのアプローチを実現する。各計測・分析においては、下記のフェーズに分かれており、それぞれに必要なインプット/アウトプットのデータ与件を確認・整理する。

成分分析：前処理、成分分析

シーケンシング：一次解析としてのライブラリー調整(試薬/調整手法)、シーケンシング

イメージング：前処理、分注、イメージング

各フェーズにおいて、シミュレータにフィードバックする一連のデータ生成を確認すると共に、バイオデータ連携基盤に接続する際のデータスキーマやフォーマットの要件を抽出する。加えて、モデル植物や培養細胞のメタデータ、解析回数なども参照できるようにしておき、モデル植物や実験計画の変更に伴い、どの範囲までチューナブルなものにするか、検証する。

## 3) 研究結果

■実験・計測環境を活用した成分分析、シーケンシング、イメージングから得られたデータのシミュレーションとの比較、及び、抽出したデータ要件

①LC/MSを対象に、遊離アミノ酸を分析したデータの構造を確認。

LC/MSによる解析データは、30を越えるアミノ酸の溶出時間、ピーク面積、濃度などがアノテーションされて出力されるものと、質量分析計データ互換フォーマット「mzXML」形式で成分（ピーク）ごとに出力されるものがある。

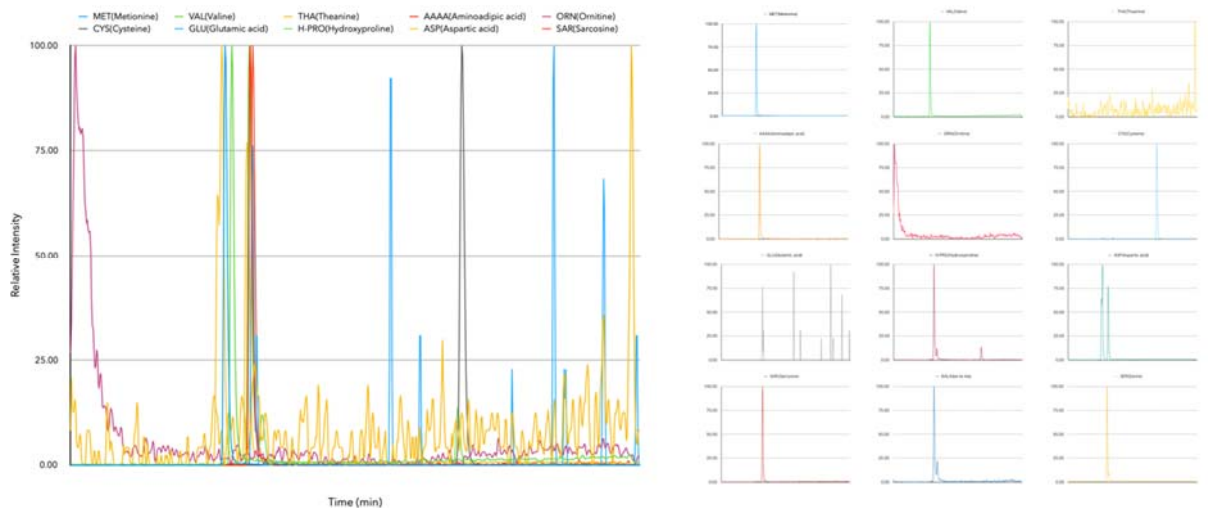


図1 遊離アミノ酸のクロマトグラムの一例

②シミュレーション結果や各数理モデルを与件とした、合成生物学実験/培養

シミュレータで単糖類の生産制御過程（単糖の生産比率にフォーカス）に関する仮説を設定することを想定し、実験を試行的に実施した。先行研究においても経路が自明な多糖は少ないが、単一の単糖から構成される多糖は候補が多い。そのためシミュレーション上、多糖を構成する単糖の構成比率を変えることで、合成過程をシミュレートすることとした。そこで、バイオラボ（パイロット版）では、上記シミュレーションが実現可能な合成生物実験/培養を試行し、バイオデータ生産拠点として高度な研究に資する良好な結果が得られた。

③上記を実現する環境構築

バイオラボ（パイロット版）において仮説検証を実施可能な環境を構築した。従来よりも高度な解析が可能だけでなく、一部ロボットによる自動化により作業の効率化が実現できる環境とした。また、仮説検証を円滑に行うため、各機器をより最適に配置した。一般的な実験/培養・測定の流れは、サンプル調整→分注→セルソート→培養→シーケンシング及び/または光学的測定、となる。この流れを円滑にすることにより、バイオラボ（パイロット版）での作業がより効率化し、仮説検証のためのバイオデータの生産性が向上することが期待できる。従来のバイオデータ生産拠点では、各機器が単体で置かれていたり、別々の場所に置かれていたり、作業の流れを阻害する物品が配置されていたりするため、本研究では、上記の一般的な実験/培養・測定の流れに沿ったレイアウトとした。





図2 バイオラボ（パイロット版）の概観

#### ④出力データの構造確認

現状、データフォーマットには機器メーカーごとに独自のものや、共通するものが存在するが、それらのバリエーションを勘案したバイオデータ連携基盤への接続を実現するため、各機器の出力データの構造を確認した。概して、1つの出力ファイルにはヘッダ部とデータ部があり、ヘッダ部に主に設定情報が含まれる。

また、各機器からの出力データの他に、研究において不可欠な情報があることが明らかとなった。すなわち、各機器で解析を実行する前処理に係るプロトコルや試薬、培養条件等の情報である。これらの前処理に関する情報は、研究の核となる部分であり、研究者によって各研究テーマに応じて操作され（試薬の選択・調整、培養条件の設定等）、各機器で解析されることにより、実験結果として出力される。そのため、これらの前処理等に関する情報を、実験結果（各機器の出力データ）に付随するメタ情報として確認・整理し、バイオデータ連携基盤に接続可能とした。

さらに、実験結果（各機器の出力データ）は単に文字列や数値の羅列であることが多く、それ自身を見ても情報としての価値がない。例えば、シーケンサーからの出力である塩基配列はATCGの4文字の厩大な羅列であり、それ自身は意味を持たない。そのような実験結果（各機器の出力データ）は、研究者によって解釈され（二次解析という）、文字列や数値に意味が付与（アノテーション）される。例えば、成分分析では、実験結果（各機器の出力データ）はクロマトグラムというスペクトル波形で表現されるが、そのピークが何を示しているかは二次解析によって研究者（すなわち人間）によって解釈され、ピークに応じた成分名が付与（アノテーション）される。そのため、実験データ（各機器の出力データ/一次解析）に付随するアノテーション付の二次解析データのデータ構造も確認・整理し、バイオデータ連携基盤に接続可能とした。

#### ■バイオデータ連携基盤への接続

- ・バイオラボ（パイロット版）にて実施される、モデル植物からの細胞培養、それらを対象とした多種多様な成分分析、シーケンシング、イメージングの結果を一貫して管理することができるバイオデータ連携基盤への接続APIを構築した。

- ・関連情報を体系的に管理するため、モデル植物、培養、シーケンシング、成分分析、イメージングを論理的な管理IDを付与する事により、モデル植物を起点に各種培養および各種解析に関する情報を統括管理する。
- ・各種メタデータについては、以下の標準化情報との連携を意識しておくものとする。

撮影画像：OME OIMERO

理研（和光）Engineering Network

化合物等：NCBI

メタボローム：理研（横浜）Plant Metabolome MetaDatabase

シーケンシング：理研（筑波）RBO

統合化：理研（神戸）Open Life Science Platform

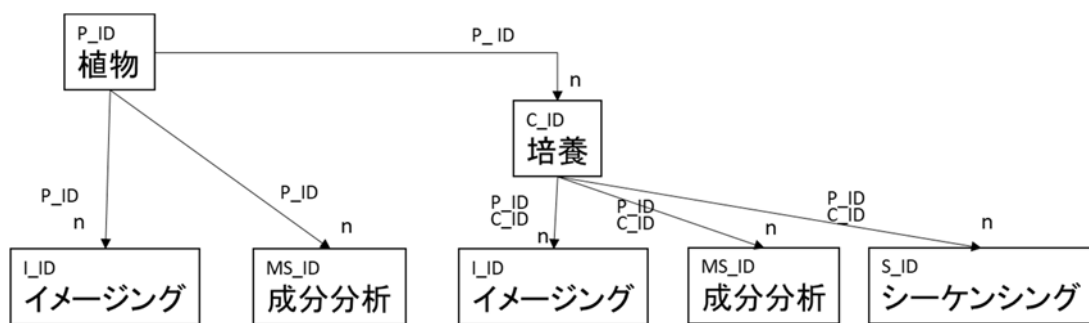


図3 バイオデータ連携基盤上でのバイオラボ関連情報の管理体系

- ・モデル植物については、多様な植物を国際標準との連携を考慮し構造化した。学名（ICBN）、科名、属名と和名、の他パスウェイシミュレータとの連携で必要となる生物種コード（NCBI（Taxon identifier））に対応し、提供される部位サンプル単位での管理を可能とした。
- ・培養については、パスウェイシミュレータとの連携を考慮し、必要情報を調整の上構造化した。
- ・成分分析については、複数の分析手法に関する、前処理、装置構成と設定条件に柔軟に対応できる構造とし、分析結果の登録はパスウェイシミュレータでの利便性及び基盤の容量制約を考慮した。
- ・シーケンシングについては、MiSeqによるDNA解析の処理内容及びシーケンシング結果情報、および解析結果として検出された遺伝子ID、遺伝子名についてパスウェイシミュレータでの利用及び基盤の容量制約を考慮し構造化した。処理に関するメタ情報としては、ライブラリー調整、サンプルシート（実行条件、データ解析条件、マニフェスト、ゲノム）の登録へ対応した。
- ・イメージングについては、前処理、撮影機器、撮影条件、撮影画像解析結果（各ウェル）を構造化した。また、画像解析の根拠となるオリジナル画像については、大容量ファイルとなるため、ファイルを特定できる情報のみを記録する。
- ・撮影機器であるOperettaCLSは、OME OMER0の仕様に基づいており、撮影に関するメタデータの一部を画像ファイルのヘッダ情報から取得することが可能である事

を確認した。

上記の成果から、パスウェイシミュレータの結果と突き合わせて、シミュレータの高度化に資するデータを生成できることを確認し、加えて、バイオデータ連携基盤とAPI接続の際の各種データ要件を明らかにし、今後、別の植物のモデルケースを設定しても、柔軟にデータ生産拠点として機能し得ることが確認できた。

#### 4) 成果活用における留意点

下記の通り、順不同で留意点を列挙する。

- ・研究目的や実験計画によって、実験プロトコル/実験回数は様々であり、その設定自体はバイオデータ連携基盤やバイオラボ（パイロット版）の活用を検討する研究者が行う必要がある。
- ・上記に関連して、全てのシミュレーションやバイオデータ連携基盤からのリクエストに対応できる訳ではなく、また全ての工程が自動化されている環境ではないため、一定のSLAが存在する。
- ・バイオデータ連携基盤に、バイオラボでの実験、解析結果を登録する際には、各研究/実験のプロトコルを踏まえ、バイオデータ連携基盤でメタ情報として登録される各種情報（例えば、ライブラリーの調整で使用する試薬や培養条件等）を研究者等が整理する必要がある。
- ・バイオデータ連携基盤に、バイオラボでの実験、解析結果を登録する際には、複数カ所からの並行登録を踏まえ各情報の管理IDの付与方法をルール化し、自動採番を行う仕組みを構築の上、各種情報構造に合わせたクライアントツールを構築することを推奨する。
- ・バイオデータ連携基盤へ登録された各種バイオラボデータを効率的に利用するためには、利用のユースケースシナリオに準じ、各APIを統括するAPI又はクライアントツールを構築する必要がある。

#### 5) 今後の課題

民間の育種を加速させるために、当該バイオラボ（パイロット版）で出来ることについてバイオデータ連携基盤などを通じて、明示的に民間企業に周知させる必要がある。具体的な研究開発での利用シーンに応じた課題や、研究推進における利便性等を本研究の成果として公開し、初めてバイオデータ連携基盤やバイオラボ（パイロット版）に関わる人でも、しっかりと当該拠点で出来ること/出来ないことを理解できる状況を作り出す必要がある。

また、本事業終了後は、シミュレータとバイオラボ（パイロット版）を組み合わせ、理研とAOIにて経過措置として運営するが、同時に、データ駆動型バイオ推進コンソーシアムへの民間企業参画を加速させ、当該コンソーシアムからの委託に基づき運営費用をまかない、且つ、特定の営利企業等の影響を必要以上に受けることなく、中立的な立場での継続的な運営を行うことが必要である。参画する民間企業が増えれば、ユースケースやデータ自体が増えることに繋がり、シミュレータ自体も高度化していくので、様々な分野においてメインとなるステークホルダーに参画を依頼し、リードステークホルダーの参画を受け

て業界全体の参画加速を行っていく必要がある。

加えて、様々な育種に係る実験やシミュレーションに対応するため、今回構築したバイオラボ（パイロット版）を参考とすることで、別の研究機関・研究場所でも複数のデータ生産拠点が構築されることが望ましく、バイオデータ連携基盤が、これら複数のバイオラボと柔軟に接続することにより、幅広い研究への対応と育種ターゲットの加速を図る必要がある。

#### <引用文献>

- Harper AD, Bar-Peled M. Biosynthesis of UDP-xylose. Cloning and characterization of a novel Arabidopsis gene family, UXS, encoding soluble and putative membrane-bound UDP-glucuronic acid decarboxylase isoforms. *Plant Physiol.* 2002;130(4):2188-2198. doi:10.1104/pp.009654
- 石水 毅, 植物細胞壁多糖キシランの生合成機構解明へ向けて, *Trends in Glycoscience and Glycotechnology*, 2008;20(113):171-172. doi:10.4052/tigg.20.171
- 前原 智子, 金子 哲, 【特集:  $\beta$ -結合糖質の有効活用-ヘミセルロースの利用-】 緒言: ヘミセルロースの構造と分解酵素-キシラン分解酵素を例としたヘミセルロースの分解機構について-, *応用糖質科学: 日本応用糖質科学会誌*, 2012; 2(3):165-168. doi:10.5458/bag.2.3\_165
- Ogura T, Date Y, Kikuchi J. Differences in Cellulosic Supramolecular Structure of Compositionally Similar Rice Straw Affect Biomass Metabolism by Paddy Soil Microbiota. *PLoS One.* 2013;8(6):e66919. doi:10.1371/journal.pone.0066919
- to, K., Sakata, K., Date, Y. & Kikuchi, J. Integrated Analysis of Seaweed Components during Seasonal Fluctuation by Data Mining Across Heterogeneous Chemical Measurements with Network Visualization. *Anal. Chem.* 2014;86:1098-1105. doi:10.1021/ac402869b
- Goué, N., Mortimer, J. C., Nakano, Y. et al. Secondary cell wall characterization in a BY-2 inductive system. *Plant Cell Tiss Organ Cult* 2013;115:223-232. doi:10.1007/s11240-013-0354-7
- 高部 圭司, 木材細胞壁研究: これまでの20年, これからの20年, *木材学会誌*, 2015;61(3):123-130. doi:10.2488/jwrs.61.123
- Conesa, A., Madrigal, P., Tarazona, S. et al. A survey of best practices for RNA-seq data analysis. *Genome Biol* 2016;17:13. doi:10.1186/s13059-016-0881-8 Okubo-Kurihara, E., Ohtani, M., Kurihara, Y. et al. Modification of plant cell wall structure accompanied by enhancement of saccharification efficiency using a chemical, lasalocid sodium. *Sci Rep.* 2016;6:34602. doi:10.1038/srep34602

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデータ連携基盤構築		
小課題名（実行課題名）	1 育種ターゲットの把握に関する研究開発 (2) AIを活用した演繹的アプローチに基づく、腸内マイクロバイオームのバイオメカニズムモデルの構築		
小課題 代表研究機関・研究室・研究者名	国立研究開発法人理化学研究所 光量子工学研究センター 光量子制御技術開発チーム 和田 智之		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

腸内には、まだ名前も付けられていない6000千兆個の腸内細菌が存在することが既に明らかになっているが、その解明が進んでいない。本研究グループでは、腸内細菌に関する数十年間に及ぶ取り組みを踏まえ、腸内細菌に関する約18000件のサンプルとそれから抽出した遺伝子情報、ならびにサンプル提供者の食生活、健康データを蓄積してきた。

本取り組みでは、これらデータを用い、サンプルから得られた遺伝子情報に基づく細菌叢の特徴と個人特性、生活特性、健康状態に関する相関的解析結果を利用可能とするバイオメカニズムモデルを構築する。

### 2) 研究方法

腸内細菌に関する約18,000件のサンプルから抽出した遺伝子情報を解析した細菌叢情報、ならびにサンプル提供者の個人特性、食生活、健康データを統合しバイオデータ連携基盤上に登録する方法を検討、APIを構築することで、AIなど各種分析に利用できるようにする。

また、それら情報に基づき実施される多面的な相関的解析結果から得られる知見を体系的に統合化し、特定の目的の下で活用できる仕組みを検討、バイオデータ連携基盤のAPIを構築し、一部の知見を抽出、登録することで、バイオメカニズムモデルを実現する。

### 3) 研究結果

腸内細菌に関する約18,000件のサンプルから抽出した遺伝子情報を解析した細菌叢情報、ならびにサンプル提供者の個人特性、食生活、健康データをAIなど各種分析に利用できる形式でバイオデータ連携基盤上に登録するAPIを構築し、検証した。

- ・個人サンプルから解析した細菌叢情報、ならびにサンプル提供者の個人特性、食生活、健康データとそれらの根拠となる調査項目情報をメタ情報として統合的に管理可能な形式でバイオデータ連携基盤のAPIを構築した。
- ・約18000件のサンプルデータをバイオデータ連携基盤上に構築したAPIを利用して登

録可能な形式に変換し、大量データを効率的に登録するクライアントツールを構築し、基盤への登録を行った。

・バイオデータ連携基盤のAPIを利用し、目的に合わせた各種条件にてサンプルデータの絞り込み抽出し、各種分析などに活用できることを検証した。

サンプルデータに対する各種相関的解析結果から得られる知見を体系的に整理統合化し、特定の目的に対し活用できる論理ネットワークとして機能する仕組みを検討、バイオデータ連携基盤のAPIを構築した。

・約18000件のサンプルデータに対する各種相関的解析結果などから得られる様々な知見を、その根拠となる対象の背景、範囲、それが成立する条件などと共に登録可能なバイオデータ連携基盤のAPIを構築した。

・これまでの研究成果から主要な知見を抽出し論理単位（1対のペア）に分解し、その根拠情報、成立条件と共にバイオデータ連携基盤のAPIを利用して登録、論理的なバイオメカニズムモデルとして構成、機能することを確認した。

・また、知見の論理単位の構成としては、対象と操作・変化の他、それを意味的に分類する情報を含むことで、特定の目的に合わせた探索を可能とした。

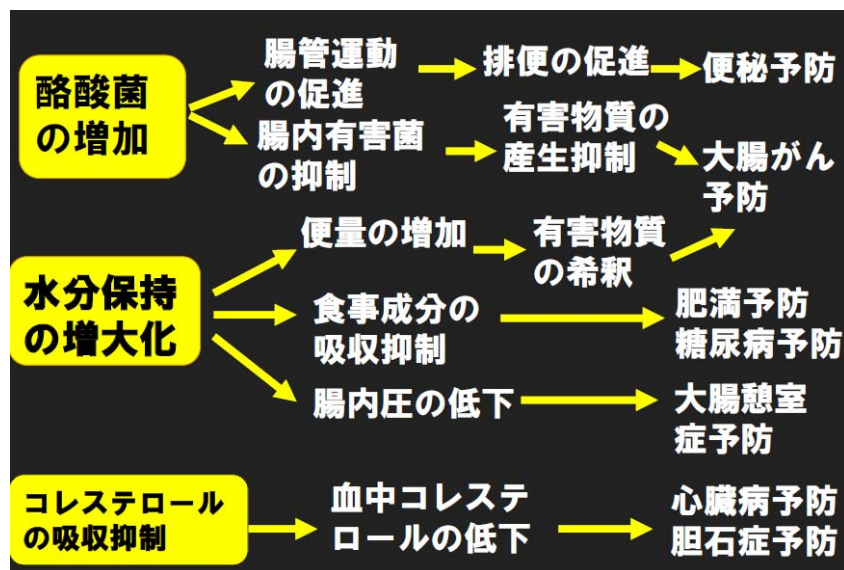


図 食物繊維の機能として搭載された論理ネットワークの例（辨野特別研究室）

#### 4) 成果活用における留意点

バイオデータ連携基盤に登録されたサンプルデータの活用にあたっては、調査時期により調査項目の改定が発生している場合があり、抽出、分析の際には関連するメタデータを参照の上、データクレンジングを行う必要がある。

バイオメカニズムモデルに対する知見の登録を逐次並行的に行う場合には、結果として構成される論理ネットワークの状況を確認する必要がある。特に、知見を削除・修正する際には、それに関連する知見への影響に留意する必要がある。

#### 5) 今後の課題

バイオデータ連携基盤に登録されたサンプルデータの活用を効率化するためには、デー

タクレンジング機能を基盤側に構築し、簡易的なGUIを通し操作することができる仕組みを提供する必要がある。また、AIなど各種解析ツールを利用した知見の獲得、評価を効率化するために、クレンジングの履歴を管理、再利用できる仕組みも必要となる。

バイオメカニズムモデルの管理を効率的に行う仕組みが必要。論理ネットワークを表示探索する仕組みや同一情報の統合化や、論理的チェック機能を提供する必要がある。

また、バイオメカニズムモデルを構成する要素として、重みづけ、論理式、数式、AI学習モデルなどを包含することで、より柔軟な知識蓄積が可能となる。

また、AIが分析結果に基づき自動的に論理モデルを構築することで、知識獲得の自動化が可能となる。

#### <引用文献>

辨野義己,「新時代を迎えた腸内常在菌研究」,日経サイエンス,2017年2月,34~43ページ  
辨野義己,「どのような個人属性・生活特性がヒト腸内常在菌構成に影響を与えるのか?」,

【特集 細菌叢解析の光と影 III.腸内細菌】,生体の科学,公益財団法人金原一郎記念医療振興財団/医学書院,2017 ; 68(2):127-131

mitsuharu Matsumoto, Ryoko Kibe, Tuji Aiba, Emiko Sawaki, Yasuhiro Koga and Yoshimi Benno, Cerebral low-molecular metabolites influenced by intestinal microbiota: a pilot study, *Frontiers in Systems Neuroscience*, Frontiers, April 2013; 7(9), doi:10.3389/fnsys.2013.00009

渡部侑子,「腸内糖代謝と腸内細菌」,腸内細菌学雑誌,2005;19:169-177

伊藤喜久治,「腸内フローラの構成と機能をめぐる新たな展開」,化学と生物,Vol.42, No.4,2004,258~264ページ

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデータ連携基盤構築		
小課題名（実行課題名）	1 育種ターゲットの把握に関する研究開発 （3）（2）で構築したモデルに基づくマイクロバイオームの健全化に資する有用成分の探索		
小課題 代表研究機関・研究室・研究者名	国立研究開発法人理化学研究所 光量子工学研究センター 光量子制御技術開発チーム 和田 智之		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

1（2）で構築したバイオメカニズムモデルを用い、健康に寄与する物質を増産させるマイクロバイオームの健全化に有用な成分の探索を行う。具体的には、これまで取得した食を含めたデータから菌類の成分比率と食に含まれる成分との相関性関係に基づき、マイクロバイオームの健全化に資する食成分を探索する。

なお、現状において、腸内細菌叢が作る酪酸を有用成分の候補としている。

### 2) 研究方法

バイオデータ連携基盤に構築したバイオメカニズムモデルAPIを利用し、健康に関連する要件を抽出し、それらの要件に寄与する有効成分の探索を行う。

### 3) 研究結果

バイオデータ連携基盤に構築したバイオメカニズムモデルAPIを利用し、健康に関連する要件として以下を抽出。

対象項目	作用・効果
便秘	予防（改善）
大腸がん	予防
肥満	予防
糖尿病	予防
大腸憩室症	予防
心臓病	予防
胆石症	予防
がん細胞	抑制



対象項目	作用・効果
腸粘膜	正常化
整腸作用	
大腸がんリスク	低下
肥満	抑制
発ガン	予防
アレルギー	低減
免疫能	調整
血圧	降下

次に、これらの健康要素に対し、バイオメカニズムモデルAPIを利用し、寄与する腸内細菌、栄養成分を探索し、以下の結果を得た。

腸内細菌種・栄養成分	
酪酸菌	増加
食物繊維	摂取
フィーカリバクテリウム・プラウスニツイ（酪酸産生菌）	多い
バクテロイデス・ブルガタス	多い
コプロコッカス（酪酸産生菌）	多い
バクテロイデーテス類（バクテロイデス門）	多い
プロバイオティクス	摂取

結果として、健康に寄与する栄養素として「食物繊維」「酪酸」「プロバイオティクス（乳酸菌、ビフィズス菌、納豆菌、発酵乳）」が抽出された。また、肥満予防への寄与する細菌種として、「バクテロイデス類」を得ることができた。

#### 4) 成果活用における留意点

バイオメカニズムモデルの探索に当たっては、該当する知見に該当する表現を知っておく必要がある。例えば、疾病（健康）に該当する知見を全て探索し、健康を増進する表現を全て調べた上で、個別に探索する必要がある。

また、探索の結果得られた知見については、その同一性についてその根拠情報や成立条件を考慮の上、判断する必要がある。

#### 5) 今後の課題

バイオメカニズムモデルの探索精度を上げるためには、関連する領域のオントロジーモデルを構築しそれに基づく論理ネットワークの構築が必要である。

また、探索に当たっては、探索の意図を理解し適切な根拠、成立条件を満たす範囲での適切な探索を行い、結果を統合化し提示する仕組みが必要である。また、得られた結果についても、その根拠となる論理ネットワークを検証・再探索できる仕組みがあると、より効果的なバイオメカニズムモデルの活用が可能となる。

<引用文献>

安藤朗「腸内細菌叢に影響を及ぼす因子」,「シリーズ腸内細菌叢9」,モダンメディア,  
2019;65(4):72-79

山下智也,平田健一,「腸内細菌と動脈硬化性疾患」,「シリーズ腸内細菌叢8」,モダンメデ  
ィア,2019;65(3):49-53

平山和宏,「腸内細菌叢の基礎」,「シリーズ腸内細菌叢1」,モダンメディア,  
2014;60(10):307-311

松本一政,高田敏彦,結城功勝,川上幸治,酒井隆史,野本康二,木村一雅,松本圭介,飯野久和,「新  
ガラクトオリゴ糖含有液糖がヒト腸内菌叢に及ぼす影響」,腸内細菌学雑  
誌,2004;18:25-35

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデ ータ連携基盤構築		
小課題名（実行課題名）	1 育種ターゲットの把握に関する研究開発 （4）有用成分生産に適した作物の育種を可能とするパスウ エイ型シミュレータの開発		
小課題 代表研究機関・研究室・研究者名	SyntheticGestalt 研究開発部 島田 幸輝		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

以下の3つの柱で開発を行う

#### a) Pathway Construction

活動: 与えられたタスクに必要な部分代謝経路を特定し、シミュレーションの範囲を決定するとともに、以下の2つの方法を組み合わせてシミュレータの構築を行う

- 既存データベースを用いて、代謝経路ネットワークを構築
- 与えられた実験結果を最もよく表現するように各反応の係数を予測

成果: 本研究で対象とするタスクを解くためのパスウェイ・マップ

#### b) Pathway Simulation

活動: 対象タスクに係る目的関数について、定量的予測情報の算出を行う

成果: 与えられた初期条件下で継時的な数値を計算するシミュレータ

#### c) Pathway Intervention

活動: 上記シミュレータを用いて介入オプションの決定を行う

成果: ウェット実験設計指針として、どのような遺伝的・環境的介入が良いかの仮説群

### 2) 研究方法

ある成分を生産するための植物選択と育種技術を開発するために、植物内の物質の生産過程をモデル化する。パスウェイ（植物内の物質の生産過程）を微分方程式でモデル化し、遺伝子因子・環境因子の両面から改善する方針を立て、育種にフィードバックをかける。ある関心事に対して定義される目的関数について（例えば累積収量など）、シミュレーションを通じて得られる予測値を元に、パスウェイに関して遺伝的または環境的な介入仮説を決定する。介入はウェット実験においては、遺伝子のノックアウトやノックダウン・栄養源投入などが対応し、シミュレーションにおいては、微分方程式を構成する変数や係数の追加や削除、値の固定や動的変更などで表現されるため、Digital Twin として機能する。

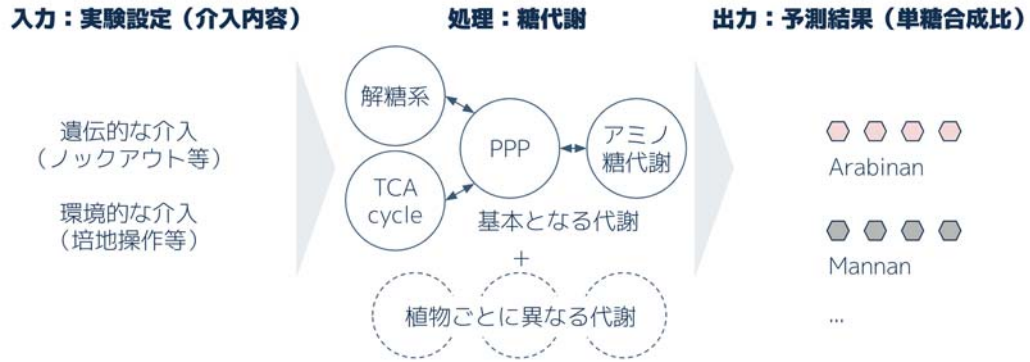


図 目的物質と対象植物に合わせたシミュレータ

### 3) 研究結果

a) Pathway Construction に必要なデータを各種データベースから入手するとともに、実験によっても実測値を取得し、データ同化を用いてパラメタ推定することで土台のパスウェイを精緻化した。

- KEGG PATHWAY
- KEGG COMPOUND
- KEGG REACTION
- KEGG ENZYME
- KEGG GENOME
- KEGG ORTHOLOGY
- SABIO-RK

上記のデータベースから、今回のシミュレーションに必要なデータをキュレーションし、シミュレータを構築するためのレシピを作成した。

#### Common Pathways

- Basic metabolism
  - [Ath00010](#) (Glycolysis: Glucose -> Pyruvate, Acetyl-CoA)
  - [Ath00020](#) (TCA Cycle: Pyruvate, Acetyl-CoA -> 2-Oxo-glutarate)
- Sub metabolism
  - [Ath00030](#) Pentose phosphate pathway
  - [Ath00040](#) Pentose and glucuronate interconversions
  - [Ath00620](#) Pyruvate metabolism
  - [Ath00760](#) (Glutamate: 2-Oxo-glutarate -> Glutamate)

#### Monosaccharide pathways

- [Ath00051](#) Fructose and mannose metabolism
- [Ath00052](#) Galactose metabolism
- [Ath00500](#) Starch and sucrose metabolism
- [Ath00520](#) Amino sugar and nucleotide sugar metabolism

図 シミュレータを構築するためのパスウェイデータ一覧（レシピ）

b) パスウェイからシミュレータの自動構築アルゴリズムを用いて、学習済みのシミュレータを計算機に実行可能な形式に変換し、小課題(5)を通して、ユーザーが利用可能なよう

にデータ連携基盤上にデプロイした。

c) シミュレータという計算機上のラボを通して仮説を策定。

#### 4) 成果活用における留意点

- ・学習済みモデルは基盤を通して利用可能な方策を取る。詳細は小課題(5)で記述
- ・c) で得た仮説は関係者の間で自由に活用可能

#### 5) 今後の課題

- ・より多くの実験データと、より多くの「実験→学習→予測→実験」のサイクルを回して実運用していくこと。

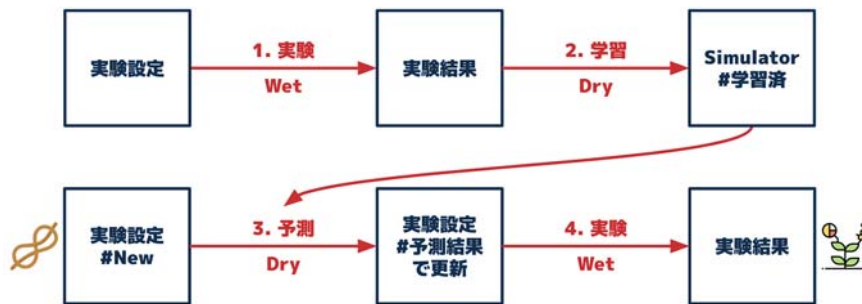


図 Digital Twin の具体的な DBTL サイクル

#### <引用文献>

細胞のシステム生物学: 江口 至洋

An Introduction to Systems Biology: Design Principles of Biological Circuits: Uri Alon

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデータ連携基盤構築		
小課題名（実行課題名）	1 育種ターゲットの把握に関する研究開発 (5) (4) で開発したシミュレーションおよび生産されたバイオデータのバイオデータ連携基盤への接続		
小課題 代表研究機関・研究室・研究者名	SyntheticGestalt 研究開発部 島田 幸輝		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

学習済みシミュレータの展開はバイオデータ連携基盤を活用して行う。その際には、データ連携基盤の特徴でもある、Dynamic API で新規追加可能な標準化されたインターフェースを全面的に利用する。

また、データのスキーマに応じて過不足を協議し、データ変換が必要であれば変換のためのマッピングの定義と提供を行う。

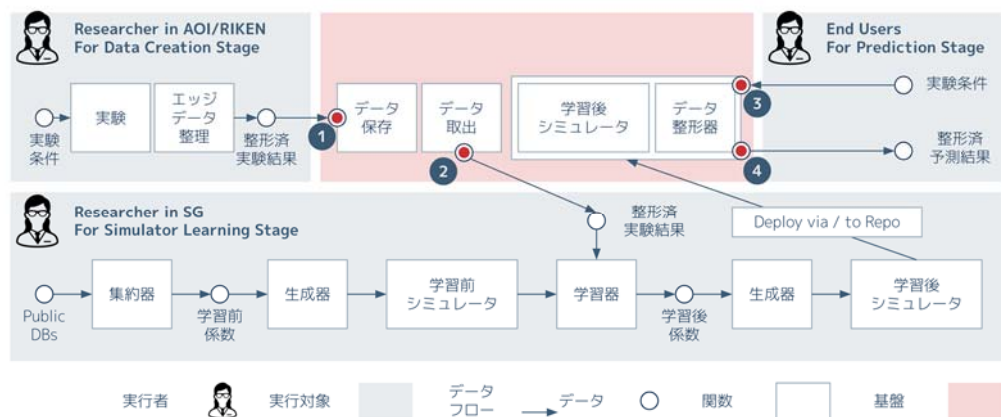


図 シミュレータに関わるシステムアーキテクチャの全体像

### 2) 研究方法

a) シミュレータのコンテンツ化：シミュレータを Data Enricher として位置付け、他のバイオインフォマティクスのツールと同様に、データに有用な情報を付加する装置とする。具体的な実装方針は、プラグイン形式にて、バイオデータ連携基盤の一部として取り込ませる。

b) シミュレータの基盤互換なスキーマの定義：アダプタは通常 ① 通信形式 (TCP/IP,

UDP, HTTP 等) ② 物理データ形式 (XML, TSV, JSON, RDF 等) ③ 論理データ形式 (項目定義ファイル) のレイヤーが存在するが、本チームでは、① HTTP ② JSON ③ 論理的に JSON で定義された微分方程式のソルバー、を指定して実施する。

項目	データ形式	遺伝子解析出力	質量分析出力	SIM 入力	SIM 出力	API ①	API ②	API ③	API ④
遺伝子条件	json/csv	✓		✓		✓	✓	✓	
環境条件	json		✓	✓		✓	✓	✓	
代謝物濃度の時間変化	json/csv			✓	✓	✓	✓		✓

図 アーキテクチャとAPI実装のマッピング

### 3) 研究結果

a) シミュレータを基盤上で仮想的に動かすことに成功した。即ち、ユーザーは実験時と同様の入力データを用いて予測による実験結果を得られる。具体的には、パスウェイで定義される全ての化合物濃度の時系列情報が出力として返却され、故に、シームレスな形でドライとウェットの往来が可能となった。

b) シミュレータのような複雑なプログラムをデプロイすることは通常難しいが、データ連携基盤のチームと連携して、ベンダー提供型・共通パーツ化の両者にも柔軟に対応できる方法論を確立した。具体的には、学習済みシミュレータを実行ファイル化し、それをコンテナ内に格納した上で、Web/API サーバを自動ビルドする方法である。このようにマイクロサービスの技術を組み合わせることで複雑性を隠蔽し、データ連携基盤の強みであるAPIベースの開発の枠組みに載せられるようにした。

### 4) 成果活用における留意点

- ・学習済みモデルはバイナリとして必要者に提供可能。

### 5) 今後の課題

- ・学習済みモデルのサービスとしての展開という観点において、ビジネスモデル確立と付随する提供方法に関する議論や検証が必要となる。ユーザーの需要密度や汎用性に基づいて、ベンダー提供型とするか、共通パーツ化とするかを選択する。



図 ユースケースとデプロイ (インストール) 方式

### <引用文献>

なし

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデータ連携基盤構築		
小課題名（実行課題名）	2 既存データ有効活用に係る研究開発 (6)最新のクラウドやネットワーク関連技術等を活用し、多様なバイオ関連研究データを相互に連携、分析を可能とするバイオデータ連携基盤の構築		
小課題 代表研究機関・研究室・研究者名	慶應義塾大学SFC研究所 政策・メディア研究科 信朝 裕行		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

本研究開発は、SIP第I期「次世代農林水産業創造技術」において、慶應義塾が中心となり構築した、「農業データ連携基盤（WAGRI、以下WAGRI）」の成果を活用し、バイオデータ連携基盤に必要とされる新たな機能を研究開発し、これらを統合した、「バイオデータ連携基盤」をWAGRIの開発に際しての実績を有するネクストスケープ社との協力体制に基づき構築する。

### 2) 研究方法

WAGRIは、我が国初のデータプラットフォームとしての基本的な機能を有し、多様な農業関連データの連携・共有・提供を、API(Application Programming Interface)により実現したもので、2019年より、農研機構において運用され、多数のデータ連携・共有・提供実績を有する。特に、データ連携に際しては、内閣官房・農林水産省が推進する、農業データの標準化ガイドラインを前提としたデータ変換機能を実装している。

また、農機メーカーや農業法人等の非IT専門業種の方々の容易な利活用を図るため、プログラミングをする事無く、Web上の管理画面を用いてAPIを作成出来る、DynamicAPI機能も有している。

これら既存成果を踏まえ、本研究開発に際しては、

- 1) 既存機能のバイオデータ連携基盤向け改修・拡張
- 2) セキュリティ機能の向上
- 3) データ取引市場の構築

の研究開発を実施する。



### 3) 研究結果

本研究では、バイオデータ連携基盤で重要となる以下の機能を開発した。

(1) 研究者が簡単に、効率的に利用するために必要な機能の拡張

- ・バイオ関連データに関する多様なデータフォーマット（CSV、JSON等）への対応
- ・利用者が直接的にデータを取り扱うことができるインターフェースの開発（ドラッグ&ドロップによるデータ登録機能）

(2) 研究者が安心して使えるための機能開発

- ・WAGRIはBtoBtoCのビジネスモデルであり、法人（B）を利用者として想定しているため、データのアクセスコントロールも法人（企業ID）単位が前提となっていたが、バイオデータ連携基盤では、法人の中に所属している一研究者単位でのアクセスコントロールが必要になることも想定されるため、個人（個人ID）単位でアクセスコントロールできる機能を開発した。

(3) 研究データが積極的に活用されるために、以下の特徴を有するデータ取引市場機能を開発した

- ・会員種別（一般会員、プレミアム会員）により、データ活用範囲を定義できる機能
- ・利用者が希望するデータの検索性向上のためのマッチング機能（アノテーション、検索等）
- ・その他、データ取引に必要となる決済、請求機能など

### 4) 成果活用における留意点

バイオデータ連携基盤上にデータを登録する際には、基盤の利用者IDの申請と付与が必要となる。この時、データ提供者に関しては、データを保有している研究者と、そのデータを基盤上に登録・接続する担当者が異なることがあり、研究者はシステムに明るいITベンダー等に登録・接続作業を依頼することがある。この時、データ保有者と登録・接続ベンダーとの間に、適切な契約と、基盤利用に関するシステム上の適切な権限付与には注意が必要となる。

### 5) 今後の課題

今回構築したバイオデータ連携基盤および取引市場は、令和2年4月以降は、一旦、慶應義塾大学により運営されるが、今後、より適切なビジネスモデルの検討を行い、必要に応じて、半官半民の組織等への運営組織の移管も検討する。

### <引用文献>

なし

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデータ連携基盤構築		
小課題名（実行課題名）	2 既存データ有効活用に係る研究開発 (7) バイオ分野におけるデータ利用規約（案）の策定		
小課題 代表研究機関・研究室・研究者名	慶應義塾大学SFC研究所 政策・メディア研究科 信朝 裕行		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

今回構築を行うデータ連携基盤を活用して、様々な主体が円滑なデータ連携を可能とするためには、これまでの取組みを十分に踏まえながら、新たなバイオ分野におけるデータ利用規約を策定することが重要となる。

本研究では、WAGRIの運用規則として本研究グループが中心となり農林水産省と連携して作成済みのデータ利用規約を踏まえ、その他各省から発表済みのガイドラインを参照しつつ、AI等の新たなデータ活用のあり方も射程に入れた、新たなバイオ分野におけるデータ利用規約（案）を策定する。

### 2) 研究方法

バイオ分野におけるデータ利用規約の策定に当たっては、これまで経済産業省や農水省において検討が行われてきたデータ利用規約について検証し、さらにこれに加えて、以下のような視点から再検討しその素案を策定する。

- ・情報の機微性や本人同意の重要性といったバイオ分野ならではの特質を考慮
- ・データ連携基盤を通じデータおよびデータ活用が取引される可能性を反映
- ・AIを含めた活発なデータ連携と利活用のための視点を加味

中でも、AI等によるデータ利活用を想定して重要となる、ELSI観点については、先行するハーバード大学、MIT等の論文を中心としたデスクリサーチに加え、国内研究者（バイオ関連、AI関連、個人情報保護関連）および弁護士等へのインタビュー、ディスカッションを行って、主要な論点を取りまとめた。

なお、データ利用規約の検討にあたっては、骨子を提示した上で有識者会議にて内容の精査を行う予定であったが、状況に鑑みて（コロナウイルス感染拡大を避けるため）、持ち回り形式での検討にこれを替える。

### 3) 研究結果

これまでの各省ガイドラインを踏まえ、バイオ領域ならではの特質を加味し、さらに今後想定されるAI活用の観点からの要請にも十分配慮したバイオデータ利用規約（案）を策定した。また、当規約においては、データの直接的な取引のみならず、基盤を通じた分析等の利用についても対応することでデータ提供元である研究機関や研究員の懸念を低減し、実践的なデータ連携を加速しうるものとした。

#### 4) 成果活用における留意点

今回の成果の活用においては、個人に関わる情報については本人の同意を得たものを対象としていることを考慮しなければならない。本人を得ていない個人情報や要配慮個人情報の扱い、例えば匿名化情報に関わる取り扱いについては、関連の個人情報保護法やその他各府省のガイドラインに従うものとする。

#### 5) 今後の課題

個人に関わる情報以外にも、当該基盤を通じてやりとりされる情報については、ELSIの観点も含め継続的な議論が必要であり、今後の法制度の改定や各府省によるガイドラインの改定、さらには社会的な倫理等に関わる議論を踏まえて継続的に検討が不可欠。

#### <引用文献>

- ・ 医療・介護関係事業者における個人情報の適切な取扱いのためのガイダンス（平成29年、厚生労働省）
- ・ AI・データの利用に関する契約ガイドライン（平成29年、経済産業省）
- ・ 限定提供データに関する指針（平成31年、経済産業省）

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とバイオデータ連携基盤構築		
小課題名（実行課題名）	2 既存データ有効活用に係る研究開発 (8) 非活用データの連携・活用に向けたユースケース開発		
小課題 代表研究機関・研究室・研究者名	慶應義塾大学SFC研究所 政策・メディア研究科 信朝 裕行		

## II. 小課題（実行課題）ごとの研究目的等

### 1) 研究目的

「育種ターゲットの把握の加速化に係る研究開発」の研究グループと連携し、同グループが提供予定の18,000件の腸内マイクロバイオーームデータをバイオデータ連携基盤上で連携・活用可能なことを検証する。

### 2) 研究方法

実際の利用は理研と利用者との契約が前提となる。SyntheticGestalt社が（4）で開発するパスウェイ型シミュレータの結果と、バイオラボから生産されるデータとの連携を図ることによって、2箇所以上の研究拠点（AOI機構、理化学研究所を想定）において具体的な研究開発での利用シーンに応じた利活用への課題や研究推進における利便性等を検討し、ユースケースとして取りまとめる。

### 3) 研究結果

18,000件の腸内マイクロバイオーームデータの入手に興味のあるキリン社をユーザーとし、（6）で構築したデータ連携基盤およびデータ取引市場にて、データを活用するユースケースを開発した。また、キリン社へのヒアリングの結果から、対象データそのものの閲覧だけでなく、基盤上で研究者が対象データを容易に解析できる機能があると便利であるという意見を踏まえて、解析環境も導入した。

### 4) 成果活用における留意点

「2）研究方法」に記載とおり、データの利用はデータ提供者が定める提供範囲での利用が前提となるため、その範囲を超えての利用はできない点には注意が必要となる。

### 5) 今後の課題

今回のキリン社が18,000件の腸内マイクロバイオーームデータに興味を持ち、データ取引市場でのユースケース開発に至った。これは、利用者にとって以下の2点が重要であるこ

とを意味している。

[1]魅力的なデータがデータ取引市場上存在すること

[2]対象データが見つけやすいこと

よって、今後の課題としては、[1]を受けて引き続き、利用者にとって魅力的なデータを基盤上に保有するためのデータ提供者の発掘と、[2]を考慮したデータのアノテーションおよび検索性の向上が必要と考える。

#### <引用文献>

なし

中課題番号	19314849	中課題 研究期間	令和1年度
小課題番号	101	小課題 研究期間	令和1年度
中課題名	育種を加速するパスウェイ型シミュレータの開発とパイオデータ連携基盤構築		
小課題名	3 腸内マイクロバイオームを対象としたシングルセルゲノム解析システムの開発 (9) シングルセルゲノム解析システムの構築 (10) 腸内細菌の時系列解析 (11) ゲノムデータの情報連携基盤との共有		
小課題 代表研究機関・研究室・研究者名	早稲田大学・竹山春子 bitBiome株式会社・細川正人		

## II. 小課題ごとの研究目的等

### 1) 研究目的

ヒト腸内のマイクロバイオームは、宿主の健康状態に密接に関連していることが近年明らかになっており、培養法やメタゲノム法を用いた大規模なゲノムの解析例が多数報告されている。一方で、大部分の腸内マイクロバイオームの機能は未だに明らかになっておらず、マイクロバイオームの健全化を評価するためには、個々の微生物の機能をより詳細に理解する必要がある。早稲田大学の竹山春子およびbitBiomeの細川正人は、マイクロ流体デバイスにより作成するピコリットル容量の微小液滴（ドロップレット）を反応場として用いた、シングルセルゲノム解析技術の開発を行ってきた。

本研究では、これまでに開発したシングルセルゲノム解析手法(SAG-gel法)のハイスループット化を行い、腸内マイクロバイオームの時系列ゲノムデータの取得を可能にするシステムを構築することを目的とする。目的達成までのタスクとして、(1)シングルセルゲノム解析システムの構築、(2)腸内細菌の時系列解析、(3)ゲノムデータの情報連携基盤との共有、の3つを設定した。

### 2) 研究方法

#### (1)シングルセルゲノム解析システムの構築

SAG-gel法は、ピコリットル容量のゲルビーズ内に標的となる細菌をシングルセルレベルで封入した後、細胞溶解からゲノム増幅に至るまでの一連の工程を行うことで、シングルセルからの超並列な全ゲノム増幅を可能にする技術である。

本研究では、ドロップレットの高速作成装置や微量分注機などの機器を導入することにより、ドロップレットの作成速度の向上およびDNA増幅後のドロップレットを用いた次世代シーケンサー用ライブラリー調製工程のハイスループット化を行い、一般実験室で再現可能なプロトコルの確立を試みた。(早稲田)

## (2)腸内細菌の時系列解析

理化学研究所辨野先生のグループより提供された経時取得された糞便サンプル(連続5日間(計6サンプル))を用いて腸内マイクロバイオームのシングルセルゲノム解析を実施した。各サンプルから調製された細菌画分をドロップレット内に封入し、シングルセルレベルでのゲノム増幅を行った後、DNA結合性色素を用いてドロップレット内に増幅されたDNAを標識した。1サンプルにつき96個の蛍光陽性ドロップレットを、セルソーターを用いて96-wellプレートに分取し、(1)で構築した解析システムに基づいてシーケンスライブラリの作製および塩基配列の取得を行った。(bitBiome)

## (3)ゲノムデータの情報連携基盤との共有

Miseqから出力された個々の配列情報を元に、*de novo assembling*および生成されたContigの評価を行う解析パイプラインの構築を行った。また、取得された大規模な配列情報をバイオデータ連携基盤に接続し、共有した。(早稲田・bitBiome)

# 3) 研究結果

## (1)シングルセルゲノム解析システムの構築

機器の導入により、一度に384個のシングルセルを対象とした配列解析を可能とする解析システムを構築した。本解析システムの構築により、一日最大768個のシングルセルゲノム解析の並行処理が可能となり、従来までの手法に比べてスループットが10倍以上に向上した。また、作業工程の自動化を実施することにより、次世代シーケンサーから出力されるサンプル毎のデータ量のばらつきが抑制され、より均質な配列情報の取得が可能になった。

## (2)腸内細菌の時系列解析

経時的に採取した6点分の糞便サンプルから計576サンプルのシーケンスを行った。ゲノムの品質を評価するツールであるCheckMを用いて、得られたゲノム情報の完全性(Completeness)および冗長率(Contamination)を評価した(図1)。この結果、17科、49属、56種の細菌からなるシングルセル由来のゲノム情報(SAG: Single amplified genome)の取得が確認された。得られたゲノム情報は、主要な腸内細菌として知られるバクテロイデス門、ファーミキューテス門の細菌の情報を含んでおり、取得日によらず安定した菌叢を呈することが明らかとなった(図2)。

シングルセル由来のゲノムの平均長は $1.03 \pm 0.31$  百万塩基であり、82%のゲノムが1000 塩基以上の16S rRNA遺伝子情報を有していることが明らかとなった。従来の細菌叢解析では、400塩基程度の16S rRNA遺伝子を対象に種組成を評価するが、本シングルセルゲノム解析技術を用いることで、読み取り塩基長は平均2500倍になった。ゲノム上にある多様な遺伝子配列を調べることで、各細菌の機能的評価が可能であるため従来評価できなかった各腸内細菌の役割を知ることができる。

また、NCBIに登録された既知細菌のゲノムデータベースを参照した結果、28種の細菌がデータベースに記載のない新種に相当することが判明した。このように、本研究において開発した解析システムを腸内マイクロバイオームの解析に適用することにより、未記載の細菌を含む大量のシングルセルゲノム情報が取得できることが実証された。

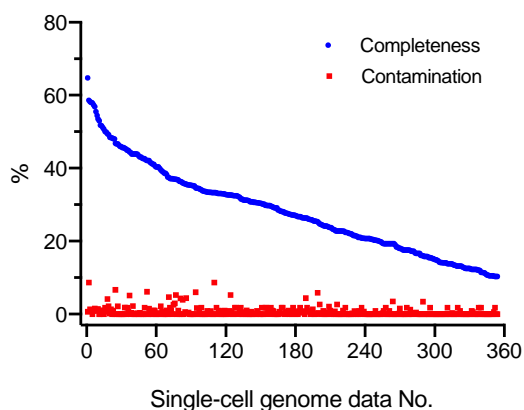


図1 腸内細菌ゲノムの完全性と冗長率

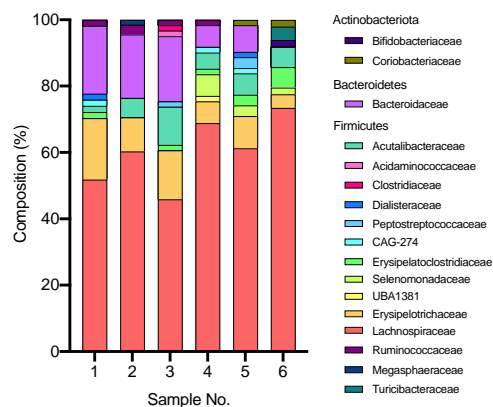


図2 腸内細菌叢の経時変化

(シングルセル由来のドラフトゲノム情報から評価)

### (3)ゲノムデータの情報連携基盤との共有

(2)にて得られた配列情報を情報連携基盤に共有した。また、得られたゲノムの平均長、GC%、データベース上での最近縁種とその相同性などの情報をまとめたファイルを出力し、必要に応じて参照できるシステムを構築した。

### 4) 成果活用における留意点

特になし

### 5) 今後の課題

本研究では、一人の被験者を対象にした経時的な糞便サンプルの採取およびシングルセルゲノム解析を行った。本研究で構築した解析システムにより、糞便中に含まれる細菌のシングルセルゲノム情報を全遺伝子レベルで比較解析することが可能になった。

今回の解析では、計測期間中に菌叢の大きな変動は確認されなかったが、今後、被験者数およびサンプルの取得日数を増やしてサンプル数を増やしていく事により、ホスト毎の細菌のゲノム情報の違いや、長期間に渡る細菌の定着や変動を種レベル・株レベルで詳細に解析することができると考えられる。

### <引用文献>

なし



### Ⅲ 研究成果一覧【公表可】

課題番号 1.9E+07

中課題名 育種を加速する・パスウェイ型シミュレータの開発と・  
バイオデータ連携基盤構築

#### 成果等の集計数

課題番号	学術論文		学会等発表(口頭またはポスター)		出版図書	国内特許権等		国際特許権等		PCT	報道件数	普及しうる成果	発表会の主催(シンポジウム・セミナー)	アウトリーチ活動
	和文	欧文	国内	国際		出願	取得	出願	取得					
19314849	0	0	4	0	0	0	0	0	0	0	3	0	1	0

注1)「学術論文」や「学会等発表」等の件数は直接本事業の成果を含むものに限定してカウントすること。

注2)特許権等のカウントは直接本事業の研究成果によるものに限定すること。

注3)特許権等について、出願公開前で知財マネジメント上、公表することが望ましくない場合に限り、「非公表」の様式に計上し記入すること。

(以下、(1)～(9)に上記集計表の明細を記載する。)

#### (1)学術論文

区分:①原著論文、②その他論文

整理番号	区分	タイトル	著者	機関名	掲載誌	掲載論文のDOI	発行年	発行月	巻(号)	掲載ページ
1		該当なし								

注1)和文、欧文の順で記載。発行年は発行年月(西暦年以下同じ)とする。

注2)区分①の原著論文(受理されたものに限る)は、謝辞等に本事業予算の支援を受けたことが明記されていること。また、論文は直接本事業の成果を掲載したものに限定して記載すること。

注3)「機関名」は当該成果に関与した代表・共同機関名を記載する。

注4)論文のDOI(Digital Object Identifier)を登録している場合は、そのDOIを記載する。登録がない場合は「なし」と記載する。

#### (2)学会等発表(口頭またはポスター)

整理番号	タイトル	発表者名	機関名	学会等名	発行年	発行月
1	パスウェイ型シミュレータについて	厚見悠	SyntheticGestalt株式会社	「育種を加速するパスウェイ型シミュレータの開発とバイオ連携基盤構築」シンポジウム	2020	3
2	シングルセルゲノム解析システムについて	竹山春子	早稲田大学	「育種を加速するパスウェイ型シミュレータの開発とバイオ連携基盤構築」シンポジウム	2020	3
3	バイオデータ連携基盤について	神成淳司	慶應義塾大学	「育種を加速するパスウェイ型シミュレータの開発とバイオ連携基盤構築」シンポジウム	2020	3
4	シミュレータおよびデータ連携基盤の社会実装について	前田浩希	株式会社電通	「育種を加速するパスウェイ型シミュレータの開発とバイオ連携基盤構築」シンポジウム	2020	3

注1)「機関名」は当該成果に関与した代表・共同機関名を記載する。

(3) 出版図書

区分: ①出版著書、②雑誌(学術論文に記載したものを除く、重複記載をしない。)、③年報、④広報誌、⑤その他

整理番号	区分	著書名(タイトル)	著者名	機関名	出版社	発行年	発行月
1		該当なし					

注1) 機関名は当該成果に関与した代表・共同機関名を記載する。

注2) 複数機関ある場合は著者名の順番と合わせる。

(4) 国内特許権等

区分: ①育成者権、②特許権、③実用新案権、④意匠権、⑤回路配置利用権

整理番号	区分	特許権等の名称	発明者	権利者(出願人等)	機関名	出願番号	出願年月日	取得年月日
1		該当なし						

注1) 複数の機関による共同出願の場合は、主となる出願人の下に行を追加し、共同出願人の情報を記載する。

(5) 国際特許権等

区分: ①育成者権、②特許権、③実用新案権、④意匠権、⑤回路配置利用権

整理番号	区分	特許権等の名称	発明者	権利者(出願人等)	機関名	出願番号	出願年月日	取得年月日	出願国
1		該当なし							

注1) 複数の機関による共同出願の場合は、主となる出願人の下に行を追加し、共同出願人の情報を記載する。

注2) 特許協力条約(PCT: Patent Cooperation Treaty)に基づく出願の場合は、出願国に「PCT」と記載し、当該様式冒頭の「成果等の集計数」欄には1件として記載する。

## (6) 報道等

区分: ①プレスリリース、②新聞記事、③テレビ放映、④その他

整理番号	区分	記事等の名称	機関名	掲載紙・放送社名等	掲載年月日	備考
1	①	シンポジウム実施のお知らせ(プレスリリースへのリンク)	株式会社電通	JAMJAM!	2020/3/6	<a href="https://iamiam.dentsukvoto.com/topics/article/26">https://iamiam.dentsukvoto.com/topics/article/26</a>
2	①	シンポジウム実施のお知らせ(プレスリリースへのリンク)	株式会社電通	Smartcell & Design	2020/3/6	<a href="https://www.smartcell.design/information/news/2020/03/473">https://www.smartcell.design/information/news/2020/03/473</a>
3	①	シンポジウム実施のお知らせ(プレスリリースへのリンク)	株式会社電通	Dentsu food (x)	2020/3/6	<a href="https://dentsufoodx.com/news/59/">https://dentsufoodx.com/news/59/</a>

注1)「機関名」は当該成果に関与した代表・共同機関名を記載する。

注2)「掲載誌、放送社名等」には同様の記事が複数社で報道された場合は全ての社名を記載する。

注3)Web上に掲載している場合は、「備考」にURL等を記載すること。

## (7) 普及に移しうる成果

区分: ①普及に移されたもの・製品化して普及できるもの、②普及のめどがたったもの、製品化して普及のめどがたったもの、③主要成果として外部評価を受けたもの(複数選択可)

整理番号	区分	成果の名称	機関名	普及(製品化)年月	主な利用場面	普及状況
1		該当なし				

注1)機関名は当該成果に関与した代表・共同機関名を記載する。

(8) 発表会の主催(シンポジウム・セミナー等)の状況

整理番号	発表会の名称	機関名	開催場所	年月日	参加者数	備考
1	「育種を加速するパスウェイ型シミュレータの開発とバイオ連携基盤構築」シンポジウム		ネット開催	2020/3/18	100	新型コロナウイルス対策のためネット開催に変更した。

注1)機関名は当該成果に関与した代表・共同機関名を記載する。

注2)概要等をWeb上に掲載している場合は、「備考」にURL等を記載すること。

(9) アウトリーチ活動の状況

区分: ①一般市民向けのシンポジウム・講演会及び公開講座・サイエンスカフェ等、②展示会及びフェアへの出展・大学及び研究所等の一般公開への参画、③その他(子供向け)

整理番号	区分	アウトリーチ活動	機関名	開催場所	年月日	参加者数	主な参加者	備考
1	①	該当なし						

注1)機関名は当該成果に関与した代表・共同機関名を記載する。

注2)概要等をWeb上に掲載している場合は、「備考」にURL等を記載すること。