

令和 6 年 5 月 27 日現在

機関番号：12102

研究種目：挑戦的研究（萌芽）

研究期間：2021～2023

課題番号：21K19815

研究課題名（和文）細胞培養のデータ駆動型最適化の基盤技術の開発

研究課題名（英文）Development of fundamental technologies for data-driven optimisation of cell culture

研究代表者

イン ベイウエン（YING, Bei-Wen）

筑波大学・生命環境系・准教授

研究者番号：90422401

交付決定額（研究期間全体）：（直接経費） 4,900,000円

研究成果の概要（和文）：本研究は個人の感覚に頼る細胞培養技術に、データサイエンスのアプローチを取り入れることにより、細胞培養の再現性や生産性といった問題の解決を目指した。網羅的実験により、培養条件と培養の良し悪しが繋ぐビッグデータを蓄積し、世界初となる「環境情報-細胞動態」のデータベースを構築した。実験データに機械学習を適用し、どの環境因子がどの程度に細胞培養に寄与するのかを予測し、それらの相対重要度を明らかにした。アクティブラーニング（実験と学習の繰り返し）により、学習モデルの予測精度を向上させ、細胞培養の培地の最適化に成功した。本研究で確立されている培養実験と機械学習の方法が様々な細胞培養に活かせることになる。

研究成果の学術的意義や社会的意義

細胞培養は基礎研究においても、健康と医療のための産業応用においても、普遍的技術であるにも拘わらず、細胞培養の再現性、安定性、安全性の課題が長年に渡って未解決のままである。本研究はアナログ方式で行われている細胞培養に、デジタル方式である応用情報科学的研究手法を取り入れることにより、細胞増殖に対する高度な制御を実現し、細胞培養の諸問題を解決することに繋がる。本研究は、機械学習・人工知能が職人芸的な細胞培養にある根幹問題を解決する初めての挑戦であり、その成功が生命科学研究において抜本的なトレンドチェンジとなる。本研究で検証される方法論が細胞培養の不確実性を減らし、再生医療の産業化に寄与する。

研究成果の概要（英文）：This study was to address issues pertaining to the reproducibility and productivity of cell culture. A data science approach was employed to improve cell culture, which is commonly based on individual senses. Through a series of comprehensive experiments, a substantial amount of data was gathered, linking specific culture conditions with the resulting quality of the culture. The dataset was then used to construct the database of 'environmental information- cell growth'. Machine learning was applied to the experimental data in order to predict which environmental factors contribute to cell culture and to what extent, as well as to identify their relative importance. Active learning (repeated experiments and learning) was employed to enhance the prediction accuracy of the learning model and optimise the culture medium for cell culture. The culture experiments and machine learning methods established in this study can be applied to a variety of cell cultures.

研究分野：システム生物学

キーワード：細胞培養 機械学習 培地 データサイエンス

1. 研究開始当初の背景

細胞の振る舞いを厳密に制御するためには、内部遺伝情報と外部環境情報の両面を把握することが不可欠である。しかし、遺伝情報が一塩基レベルまで明らかとなっており、膨大な数のデータベースが整備されているのに対し、その遺伝情報を表す細胞がさらされる環境情報がデータ化(数値化)されていない。そこで、応募者が環境情報に着目し、細胞の振る舞いと環境因子の対応関係を定量的かつ網羅的に評価することにより、培養条件と培養の良し悪しが繋ぐデータが蓄積され、世界初となる「環境情報 - 細胞動態」のデータベースが構築される。このビッグデータを持続的に充実させることにより、細胞増殖の予測と細胞培養の制御が実現される。

遺伝情報も外部環境も高度な複雑系であるため、特定の生命科学メカニズムや数理モデルによる予測が不可能である。そこで、応募者が既存知識や経験則による先入観を持たずに、網羅的実験で得られたビッグデータに対して、機械学習を適用し、細胞培養のデータ駆動型最適化の基盤技術を開発する。詳細なメカニズムが分かっているなくても、膨大な培養データに対する数値計算を自動的かつ高速に繰り返して適用することで、細胞培養の良し悪しに対して各環境因子(例えば、培地成分)の貢献度や至適条件を明らかにする。実験と学習の繰り返しにより、予測精度を向上させ、目的に応じた細胞培養の合理化を実現する。これらにより、細胞増殖がどの環境因子で決まるのかを解明することに伴い、細胞培養の再現性、安定性と安全性を保証する汎用性の高い基盤技術を確立することを目指す。

2. 研究の目的

細胞培養は基礎研究においても、健康と医療のための産業応用においても、普遍的技術であるにも拘わらず、細胞培養の再現性、安定性、安全性の課題が長年に渡って未解決のままである。そのため、高度な細胞培養が求められている再生医療では産業化の困難に直面している。そこで、本研究はアナログ方式で行われている細胞培養に、デジタル方式である応用情報科学的アプローチを取り入れることにより、生物の振る舞いに対する高度な制御に繋げ、細胞培養の諸問題を解決する。

3. 研究の方法

1) 細胞培養実験系の確立

モデル培養細胞 HeLa-S3 を対象に、培養実験の操作と細胞増殖の良し悪しを数値化することで、機械学習を可能にする。細胞培養に使用する培地は、市販品ではなく、30種類の単一化学組成から混ぜ合わせることで、200種類以上のオリジナルの培地組み合わせを作り出した。これらの培地成分の種類を選定と探索濃度勾配の決定は、細胞培養に関する論文と培地製造情報を調べ、予備実験で検討して決めた。培地成分の組み合わせ方法が微生物を用いた先行研究で確立されたアプローチで実施した。

2) ビッグデータの蓄積

培地種類や実験操作の詳細といった環境因子を網羅的に変化させ、HeLa細胞の細胞密度、生細胞数など細胞培養の良し悪しを表す指標を計測し、培地の詳細情報と細胞活性がリンクさせたビッグデータを取得した。細胞培養結果の計測手法やデータ処理の高速化を検討し、効率的に実験データを取得し、機械学習に適用できる細胞培養データセットを整備した。

3) 学習プログラムの開発

上記データセットに機械学習を適用することにより、細胞培養の分析と予測を行った。これまでに微生物細胞を対象に開発した機械学習プログラムを改変し、異なるアルゴリズムで学習・予測モデルを構築した。実験データの蓄積に伴い、学習モデルの予測精度を高めた。

4) 検証

上記実験と学習分析により、個々の培地成分がどの程度、どのようにHeLa細胞の活性に寄与するのかを明らかにした。細胞活性を向上するために、アクティブラーニング(実験と学習の繰り返し)により培地構成の最適化を行い、培養条件の学習式最適化の成否を検証した。

4. 研究成果

網羅的培養実験と計測により、培地構成の詳細とHeLa-S3細胞培養の良し悪し(細胞活性)が繋ぐデータベースが構築された。実験データに対する学習分析により、各培地成分が培養細胞の活性への寄与度が明らかとなった。アクティブラーニング(実験と学習の繰り返し)により、学習モデルの予測精度が向上され、高い活性を示す細胞培養を実現する培地構成が最適化された。スケールアップ培養実験により、機械学習による最適化された培地の有効性が検証された。これらの研究成果の詳細は下記の図表(公表した学術論文から引用)を用いて説明する。

1) 細胞培養実験系の確立とデータ取得

まず、実験するための初期細胞濃度を確定した。HeLa-S3細胞の初期細胞濃度は 10^4 cells/mlとした。これより低濃度や高濃度では培養時間が長くなり(図1A)増殖速度が低下する(図1B)ためである。つぎに、細胞培養の良し悪しを定量的に評価する計測方法を決定した。細胞培養の

定量的評価は、複数の方法、すなわち、細胞粒子の計測 (Multisizer)、細胞画像分析 (BioStudio-T)、化学反応アッセイ (CCK-8)、細胞染色および血球計算盤の計数 (Haemocytometer) を試験することにした。これらの方法を、最も信頼性が高く一般的に使用されている Haemocytometer を用いた方法と比較したところ、高い相関係数を示す Multisizer と CCK-8 を用いた方法が好ましいことがわかった (図 1C)。操作に要する時間と細胞濃度の測定範囲 (図 1D) を考慮し、光学吸光度 (A450) で細胞濃度を評価する CCK-8 を用いた細胞 NAD(P)H 量に対する化学反応測定法を選択した。この方法は、ハイスループットで実施できるため、機械学習用の広範なデータセットを取得するのに効率的で便利であった。

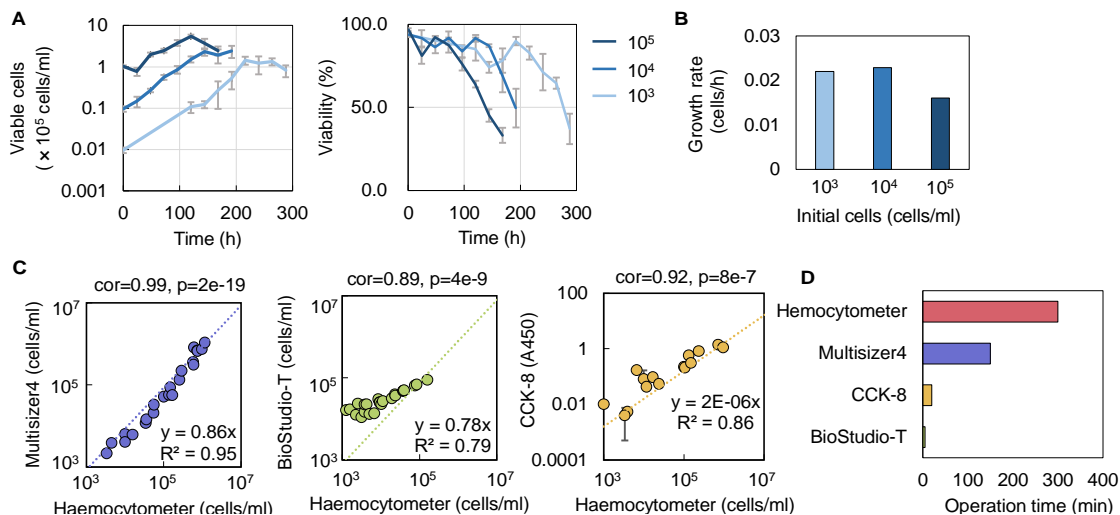


図 1 実験系の確立

A) 増殖曲線 B) 増殖速度 C) 計測方法の比較 D) 計測にかかる時間

2) アクティブラーニングによる細胞培養の培地最適化

最適化の対象となる培地成分は、一般的に使用されている Eagle's minimum essential medium (EMEM) の組成に従って決定し、約 31 成分からの構成とした。フェノールレッドとペニシリン-ストレプトマイシンを除いた 29 成分を用いて、様々な培地の組み合わせを調製した (図 2A)。これらの成分の濃度勾配は、幅広いデータバリエーションを得るために対数スケールで変化させ、生物学的測定や実験経験から偏らないようにした。機械学習は、広範囲の化学濃度を通して、従来の培地最適化では試されることのない培地の組み合わせを探索することができた。最終的に、232 種類の培地の組み合わせで細胞培養を行い、24 時間または 48 時間間隔で細胞培養の時間的変化を測定した (図 2B)。最終的に、数千の A450 の記録データが得られた。

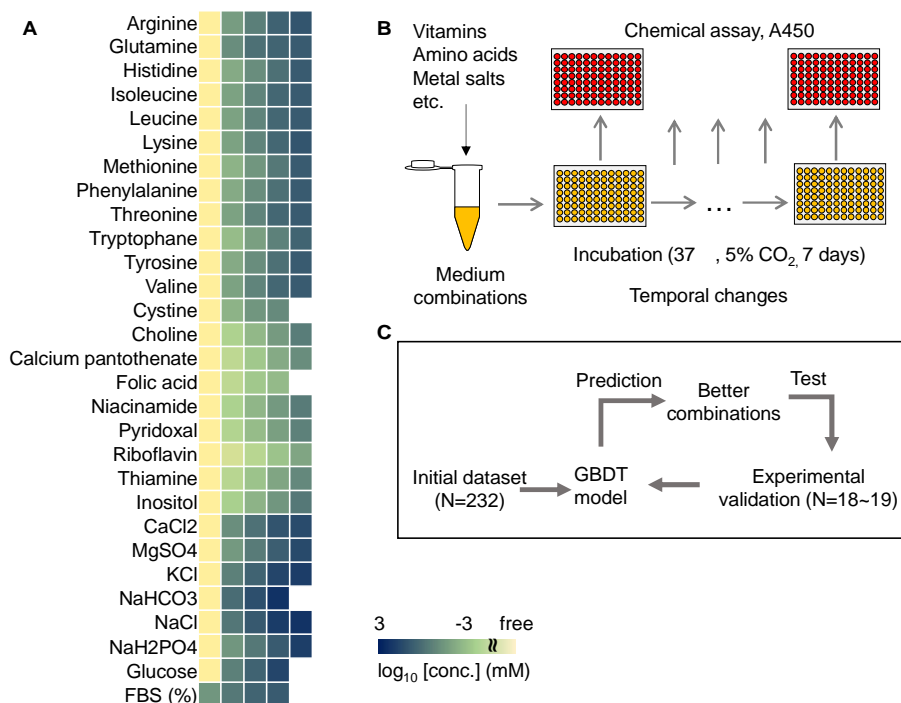


図 2 アクティブラーニングの概要

A) 培地の組合せ B) 実験手順 C) アクティブラーニング (実験と学習の繰り返し)

細胞培養を向上させる培地の組み合わせを探索するために、アクティブラーニングを行った。より良い細胞培養、すなわちより高い A450 を導く培地の組み合わせを予測するために、機械学習は GBDT モデルを用いた。232 種類培地の組み合わせからなるデータセットで学習ループが開始された (図 2C)。予測された組み合わせを実験で検証した。実験結果はトレーニングデータに加えられた。モデル構築、培地予測、実験検証、トレーニングの手順を 4 回繰り返した (図 3A)。その結果、細胞培養の A450 値と GBDT モデルの精度の両方が向上した。細胞培養はラウンド 3 で有意に改善され、ラウンド 4 以降も同程度であった (図 3A)。ラウンド 3 と 4 では、74% と 58% が EMEM より優れた細胞培養を示した。学習モデルの予測精度はラウンド数が増えるにつれて徐々に向上した (図 3B)。

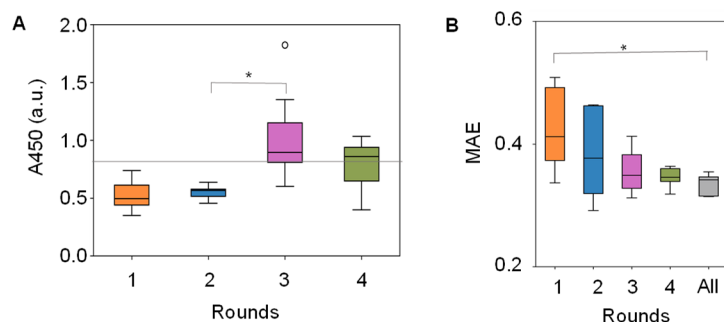


図 3 アクティブラーニングによる培地最適化

A) 細胞活性を示す A450 の上昇 B) 学習モデルの予測精度の向上

3) 細胞培養への培地成分の寄与度

培地成分の細胞培養への寄与は、GBDT を用いて推定した。全てのデータを用いて、各培地成分の重要度を学習分析した (図 4)。その結果、細胞培養を改善するために、金属塩や FBS などの類似した成分を微調整していることを示していた。興味深いことに、FBS ではなく NaCl が細胞培養を決定する主要成分であった。

4) 最適化培地とオリジナル培地との比較 (検証)

最適化された培地はオリジナル培地 EMEM よりも高い A450 を示した。スケールアップ培養実験で、予測された最適化培地と元の培地との直接比較を行った。最適化された培地は、より高い A450 を示し (図 5 左) 細胞あたりの A450 も増加した (図 5 右)。機械学習による培地最適化は、NAD(P)H の細胞内存在量をより顕著に増加させた。この発見は、細胞培養のために一般的に使用されている化学的アッセイが、細胞濃度のみを提供するものであるかどうかを再考させるものであった。

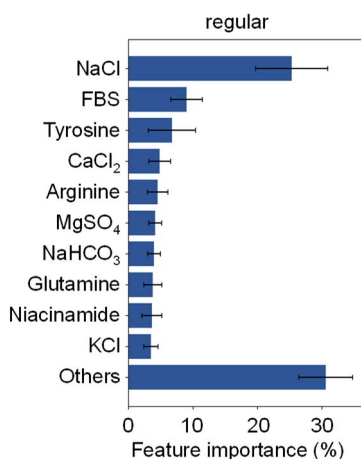


図 4 培地成分の寄与度

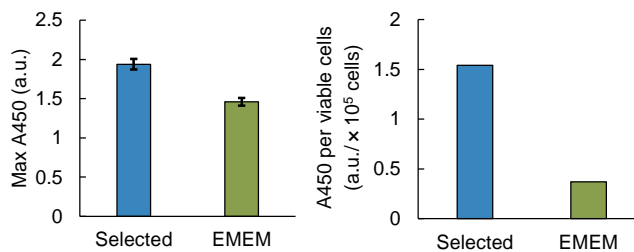


図 5 最適化培地と既存培地の性能比較

5. 主な発表論文等

〔雑誌論文〕 計12件（うち査読付論文 9件 / うち国際共著 0件 / うちオープンアクセス 11件）

1. 著者名 Hashizume Takamasa, Ying Bei-Wen	4. 巻 70
2. 論文標題 Challenges in developing cell culture media using machine learning	5. 発行年 2024年
3. 雑誌名 Biotechnology Advances	6. 最初と最後の頁 108293 ~ 108293
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.biotechadv.2023.108293	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Aida Honoka, Uchida Keisuke, Nagai Motoki, Hashizume Takamasa, Masuo Shunsuke, Takaya Naoki, Ying Bei-Wen	4. 巻 21
2. 論文標題 Machine learning-assisted medium optimization revealed the discriminated strategies for improved production of the foreign and native metabolites	5. 発行年 2023年
3. 雑誌名 Computational and Structural Biotechnology Journal	6. 最初と最後の頁 2654 ~ 2663
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.csbj.2023.04.020	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Hitomi Kenya, Ishii Yoichiro, Ying Bei-Wen	4. 巻 12
2. 論文標題 Diversity for commonality in the evolutionary changes of the reduced genome to regain the growth fitness	5. 発行年 2023年
3. 雑誌名 eLife	6. 最初と最後の頁 93520
掲載論文のDOI (デジタルオブジェクト識別子) 10.7554/eLife.93520.1	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Aida Honoka, Ying Bei-Wen	4. 巻 12
2. 論文標題 Efforts to Minimise the Bacterial Genome as a Free-Living Growing System	5. 発行年 2023年
3. 雑誌名 Biology	6. 最初と最後の頁 1170 ~ 1170
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/biology12091170	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Zhang Shuyang, Aida Honoka, Ying Bei-Wen	4. 巻 3
2. 論文標題 Employing Active Learning in Medium Optimization for Selective Bacterial Growth	5. 発行年 2023年
3. 雑誌名 Applied Microbiology	6. 最初と最後の頁 1355 ~ 1369
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/applmicrobiol3040091	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Hashizume Takamasa, Ozawa Yuki, Ying Bei-Wen	4. 巻 9
2. 論文標題 Employing active learning in the optimization of culture medium for mammalian cells	5. 発行年 2023年
3. 雑誌名 npj Systems Biology and Applications	6. 最初と最後の頁 20
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41540-023-00284-7	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 H Aida, K Uchida, M Nagai, T Hashizume, S Masuo, N Takaya, BW Ying	4. 巻 -
2. 論文標題 Machine learning-assisted discovery of growth decision elements by relating bacterial population dynamics to environmental diversity.	5. 発行年 2023年
3. 雑誌名 bioRxiv	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1101/2022.12.24.521878	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 T Hashizume, Y Ozawa, BW Ying	4. 巻 -
2. 論文標題 Employing active learning in the optimization of culture medium for mammalian cells	5. 発行年 2022年
3. 雑誌名 bioRxiv	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1101/2022.12.24.521878	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 H Aida, T Hashizume, K Ashino, BW Ying	4. 巻 11
2. 論文標題 Machine learning-assisted discovery of growth decision elements by relating bacterial population dynamics to environmental diversity.	5. 発行年 2022年
3. 雑誌名 eLife	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.7554/eLife.76846	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 K Hitomi, J Weng, BW Ying	4. 巻 13
2. 論文標題 Contribution of the genomic and nutritional differentiation to the spatial distribution of bacterial colonies.	5. 発行年 2022年
3. 雑誌名 Front Microbial	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.3389/fmicb.2022.948657	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 M Kurokawa, I Nishimura, BW Ying	4. 巻 13
2. 論文標題 Experimental evolution expands the breadth of adaptation to an environmental gradient correlated with genome reduction.	5. 発行年 2022年
3. 雑誌名 Front Microbial	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.3389/fmicb.2022.826894	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 AIDA honoka; HASHIZUME Takamasa; ASHINO Kazuha; YING BEIWEN	4. 巻 -
2. 論文標題 Machine learning-assisted discovery of growth decision elements by relating bacterial population dynamics to environmental diversity.	5. 発行年 2022年
3. 雑誌名 bioRxiv	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1101/2022.02.10.479953	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計9件（うち招待講演 2件 / うち国際学会 0件）

1. 発表者名 張抒楊、會田穂乃香、イン ベイウエン
2. 発表標題 機械学習による微生物選択培地の最適化
3. 学会等名 日本農芸化学会2023年度広島大会
4. 発表年 2023年

1. 発表者名 會田穂乃香、イン ベイウエン
2. 発表標題 遺伝情報と化学環境が大腸菌増殖に与える影響の横断的探索
3. 学会等名 日本ゲノム微生物学会第17回大会
4. 発表年 2023年

1. 発表者名 會田穂乃香、イン ベイウエン
2. 発表標題 細菌の増殖動態を決定する環境因子の機械学習支援型探索
3. 学会等名 日本農芸化学会2023年度広島大会
4. 発表年 2023年

1. 発表者名 橋詰崇雅、イン ベイウエン
2. 発表標題 細胞培養のデータ駆動式培地開発
3. 学会等名 日本動物細胞工学会JAACT2022
4. 発表年 2022年

1. 発表者名 イン ベイウエン
2. 発表標題 バイオマス生産最適化に向けたデータ駆動型システム微生物
3. 学会等名 MiCSワークショップ（招待講演）
4. 発表年 2022年

1. 発表者名 イン ベイウエン
2. 発表標題 機械学習による細胞増殖の予測とその制御
3. 学会等名 技術情報協会公開セミナー（招待講演）
4. 発表年 2022年

1. 発表者名 橋詰 崇雅、會田 穂乃香、イン ベイウエン
2. 発表標題 機械学習による微生物増殖の予測
3. 学会等名 日本ゲノム微生物学会第16回大会
4. 発表年 2022年

1. 発表者名 會田 穂乃香、橋詰 崇雅、イン ベイウエン
2. 発表標題 大腸菌増殖を決定する環境因子の網羅的探索
3. 学会等名 日本ゲノム微生物学会第16回大会
4. 発表年 2022年

1. 発表者名 會田 穂乃香、橋詰 崇雅、イン ベイウエン
2. 発表標題 微生物集団増殖に優位的に影響する環境因子の探索
3. 学会等名 日本微生物生態学会第34回大会
4. 発表年 2021年

〔図書〕 計0件

〔出願〕 計3件

産業財産権の名称 培養環境判定システム、培養装置、要求取得装置及びプログラム	発明者 イン ベイウエン、 會田 穂乃香、橋詰 崇雅	権利者 同左
産業財産権の種類、番号 特許、2022-061228	出願年 2022年	国内・外国の別 国内

産業財産権の名称 微生物培養用合成培地及び微生物の培養方法	発明者 イン ベイウエン、 會田 穂乃香	権利者 同左
産業財産権の種類、番号 特許、2021-171528	出願年 2021年	国内・外国の別 国内

産業財産権の名称 培養環境判定装置、培養システム、培養環境判定方法、培養方法、及びプログラム	発明者 イン ベイウエン、 天笠 俊之	権利者 同左
産業財産権の種類、番号 特許、2022-024033	出願年 2022年	国内・外国の別 国内

〔取得〕 計0件

〔その他〕

-研究室ホームページ https://www.u.tsukuba.ac.jp/~ying.beiwen.gf/research.html -プレスリリース https://tsukuba.flintbox.com/technologies/edab3bd3-6e71-4415-91c4-ad4e5d442bbb https://www.eurekalert.org/news-releases/990931 https://www.tsukuba.ac.jp/journal/biology-environment/20230530180000.html https://www.tsukuba.ac.jp/journal/biology-environment/20220830140000.html https://www.tsukuba.ac.jp/journal/biology-environment/20190523140032.html -ブロードキャスト https://podcasters.spotify.com/pod/show/university-of-tsukuba -YOUTUBE https://www.youtube.com/watch?v=jw-S7xi4r4Y

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------