



カイコ遺伝子発現データの拡張・公開 — 昆虫活用技術開発やデータ駆動型研究促進に期待 —

農研機構は、シルクタンパク質産生に重要な時期のカイコ幼虫においてどのような遺伝子が生体内のどこでどの程度働いているか（網羅的遺伝子発現）がわかるデータを作成・公開しました。本成果は、カイコの高いタンパク質合成能力をさらに引き出して、有用タンパク質の生産にカイコを活用する道を拓くと期待されます。

研究データ公開

カイコは絹糸を生産する家畜化された昆虫です。農研機構は絹糸の構成成分となる2種類のタンパク質（フィブロインとセリシン）が合成される器官（絹糸腺）で他の生物が持つ遺伝子を働かせて有用タンパク質をカイコに作らせる技術を以前に開発しており、これまでに動物医薬品の原薬の生産等に利用されています。現在は、この技術を改良し有用タンパク質の生産量を上げることでより広汎なタンパク質の産生に利用できるよう研究を進めています。

有用タンパク質の産生効率向上には個々の遺伝子機能だけではなく、特定条件下での遺伝子発現の変化を網羅的に捉えた遺伝子発現データが重要になります。そこでシルクタンパク質産生に重要な時期のカイコ幼虫における網羅的遺伝子発現データを取得しました。さらにヒトやショウジョウバエの遺伝子情報との対応づけを行うことでこれらのデータベースにある膨大な遺伝子情報の利用を可能にしました。得られたデータはカイコの一生で特に多様で重要な生命活動が起きる時期に得たものであるため、カイコにおける有用タンパク質産生の研究のみならず、カイコや他の生物種の基礎から応用までの様々なフェーズの研究を促進することが期待されます。そのため、本データを公開することにしました。

得られたデータは農研機構が公開しているカイコゲノムデータベース・KAIKObase（カイコベース）に公開しています。今後はこの網羅的遺伝子発現データを用いたデータ駆動型研究¹⁾を実施し、個々の遺伝子の関連性を解明して、有用タンパク質の生産を制御する遺伝子群を同定することで、カイコによる、動物やヒト用の医薬品原薬などの有用タンパク質の産生能力を向上させ、本技術の普及を目指します。

< 関連情報 >

予算：農林水産省委託プロジェクト「蚕業革命による新産業創出プロジェクト」、内閣府戦略的イノベーション創造プログラム（SIP）「昆虫(カイコ等)による有用タンパク質・新高機能素材の製造技術の開発・実用化」

問い合わせ先など

研究推進責任者：農研機構生物機能利用研究部門 所長 吉永 優

研究担当者：同 生物機能利用研究部門 昆虫利用技術研究領域

(兼) 同 基盤技術研究本部 農業情報研究センター AI研究推進室

主任研究員 横井 翔 TEL 029-838-6129

広報担当者：同 生物機能利用研究部門 研究推進部 研究推進室長 古澤 軌

TEL 070-4880-4356 プレス用e-mail nias-koho@ml.affrc.go.jp

(10月1日以降はe-mailのみの受付とさせていただきます。)

広島大学 財務・総務室 広報部広報グループ

プレス用e-mail koho@office.hiroshima-u.ac.jp

大学共同利用機関法人 情報・システム研究機構 データサイエンス共同利用基盤施設 ライフサイエンス統合データベースセンター (DBCLS) 広報担当

プレス用e-mail public_relations@dbcls.rois.ac.jp

本資料は農政クラブ、農林記者会、農業技術クラブ、筑波研究学園都市記者会、広島大学関係報道機関、文部科学省記者クラブ、科学記者会に配付しています。

※農研機構(のうけんきこう)は、国立研究開発法人 農業・食品産業技術総合研究機構のコミュニケーションネーム(通称)です。

新聞、TV等の報道でも当機構の名称としては「農研機構」のご使用をお願い申し上げます。

開発の社会的背景

カイコは絹糸生産に用いられる家畜化された昆虫です。多量の絹糸を産生する系統や高品質の絹糸を産生する系統を作製するために、農研機構は長年カイコの研究を進めてきました。一方、1980年から2000年にかけてカイコで動物やヒト用の医薬品原薬となる有用タンパク質生産を行う方法が開発されました。これにより、カイコは「有用タンパク質を作る家畜」として新たな価値が見出されました。農研機構では、絹糸の構成成分となる2種類のタンパク質（フィブロインとセリシン）が合成される器官（絹糸腺）で、外来の有用タンパク質をカイコに作らせる技術を開発しました。この技術の普及には更なる生産性の向上が求められており、それにはカイコのゲノム情報²⁾や遺伝子の発現情報等の基盤情報が重要になります。そのため2004～2009年に本研究所が中心機関となり世界に先駆けてカイコゲノム情報やそれに付随した遺伝子の発現情報の取得・整理を行い、データベース・KAIKObaseを構築し、公開しました。

研究の経緯

農研機構はKAIKObaseを通じて、カイコやその他の昆虫の様々な研究の推進に貢献してきました。一方で、ゲノム塩基配列情報を解析するシーケンス装置³⁾はこの10年で飛躍的に発展し、膨大な塩基配列データが極めて低コストで得ることができるようになりました。そこで、2019年に農研機構は東京大学、遺伝学研究所と共同でカイコの高精度ゲノム情報を取得し、KAIKObaseを大幅に改良し、公開しました。

カイコによる、有用タンパク質のさらなる産生効率の向上には、個々の遺伝子機能だけではなく、特定条件下における様々な組織における遺伝子発現の変化を網羅的に捉えた遺伝子発現データが重要になります。今回、広島大学大学院統合生命科学研究科、情報・システム研究機構データサイエンス共同利用基盤施設ライフサイエンス統合データベースセンターと共同でカイコの様々な組織における遺伝子の網羅的な発現データを取得しました(図1)。これらのデータはカイコの一生の中で特に多様で重要な生命活動が起きる時期に得たものであるため、カイコにおける有用タンパク質産生の研究のみならず、カイコや他の生物種の基礎から応用までの様々なフェーズの研究を促進することが期待されます。そのため、本データを公開することにしました。

研究の内容・意義

1. 遺伝子発現を調べた個体や組織について

絹糸（シルク）の生成に重要な時期である5齢3日目の幼虫3匹から個体別に絹糸タンパク質を生成する絹糸腺や他の主要組織を含む、合計10種類のサンプルを取り出し、RNA-seq⁴⁾データを取りました。

2. 新規遺伝子の同定及び完全長塩基配列データの取得

得られたRNA-seqデータを以前構築した高精度のゲノム配列データと組み合わせることで、これまで配列決定が困難だったセリシン遺伝子⁵⁾群の完全な塩基配列データを得ることができました。この知見によって、従来よりも機能強化された（例えば、伸縮性や強度に優れた）シルクの創出に役に立つと期待されます。

3. 信頼性が高い、網羅的遺伝子発現データの取得

全遺伝子の発現量データを上記のサンプルごとに得ることができました。データを解析した結果、異なる個体でも同じ種類のサンプルである場合は非常に発現量が似て

いることが明らかになりました。これは得られたデータの信頼性が高いことを示しています。

4. ヒトやショウジョウバエの遺伝子との対応づけ（データの連携）

ゲノム塩基配列データは基本的に4種類の塩基で構成されているので異なる生物間でも比較することができます。そこで新規に特定された遺伝子を含むカイコの遺伝子情報をショウジョウバエやヒトの遺伝子情報と対応づけました。ショウジョウバエやヒトは、これまでの研究による膨大な知見・データの蓄積がデータベースにあります。これらの知見はそれぞれの遺伝子情報に対してIDで対応づけられています。そこでカイコとショウジョウバエおよびヒトとの対応づけを行うことによって、カイコの遺伝子にショウジョウバエやヒトのデータを付けることを可能にしました（図2）。紐づけされた遺伝子の機能に関するデータを利用して、前部絹糸腺⁶⁾では生物のエネルギー源となる炭水化物の代謝に関わる遺伝子が強く働いていることが明らかになりました。これらの新規知見から、遺伝子組換え技術等によって炭水化物代謝を活性化させることで、シルクの生産性が向上する可能性を明らかにしました。

5. 全関連データの公開（誰でも再利用可能に。）

1. から4. で得られた全てのデータは公共のデータベースに公開されています。この他の解析の際の関連ファイルも全て公式のデータレポジトリに公開しており、原著論文経由で、誰でも再利用が可能です。一方、高度なデータ解析技術を持たなくてもこれらのデータを利用できるように、農研機構が公開しているカイコゲノムデータベースKAIKObase (<https://kaikobase.dna.affrc.go.jp/>)に収録し、データの検索・利用を可能にしています。KAIKObaseではキーワード検索などで遺伝子を検索し、その発現量を調べることが可能です。詳細に関してはKAIKObaseのページを参照してください。

今後の予定・期待

今回得られた遺伝子発現データは汎用性の高い情報ですので、カイコや他の生物種の基礎から応用までの様々なフェーズの研究を促進することが期待されます。

農研機構では、本研究で得られた網羅的発現量データを用いたデータ駆動型研究によって、各遺伝子間の関連性を解明し、有用タンパク質の生産を制御する遺伝子群を同定することで、従来よりも高効率に動物やヒト用の医薬品原薬などの有用タンパク質を産生するカイコの作出を目指します（図3）。

用語の解説

- 1) データ駆動型研究：事前に想定される仮説を考えずにデータを取得し、解析をすることで、仮説を導き出す方法です。
- 2) ゲノム情報：遺伝子の実体であるDNAの配列全体を指します。
- 3) シーケンス装置：DNAやRNAの必要な情報をコピーしたRNAの配列情報を決める装置です。配列情報を解析することで、どのような遺伝子を持っているか、どのような遺伝子が働いているかを明らかにできます。
- 4) RNA-seq：生体内からとりだしたRNAの塩基配列を決める手法です。この手法で生み出されたデータを解析して、取り出した生体内でどのような遺伝子が働いているか（発現しているか）を網羅的に調べることができます。

- 5) セリシン遺伝子：カイコが絹糸を作る際に作られる、糊のような役割を持つタンパク質の設計図となる遺伝子です。
- 6) 前部絹糸腺：絹糸を合成する器官である絹糸腺の一部で、絹糸タンパク質を脱水させて、絹糸タンパク質の構造を変える役割を果たしています。

発表論文

Kakeru Yokoi*, Takuya Tsubota*, Akiya Jouraku, Hideki Sezutsu and Hidemasa Bono
Reference Transcriptome Data in Silkworm Bombyx mori.

Insects 2021, 12(6), 519; <https://doi.org/10.3390/insects12060519>

*両著者同等貢献

参考図

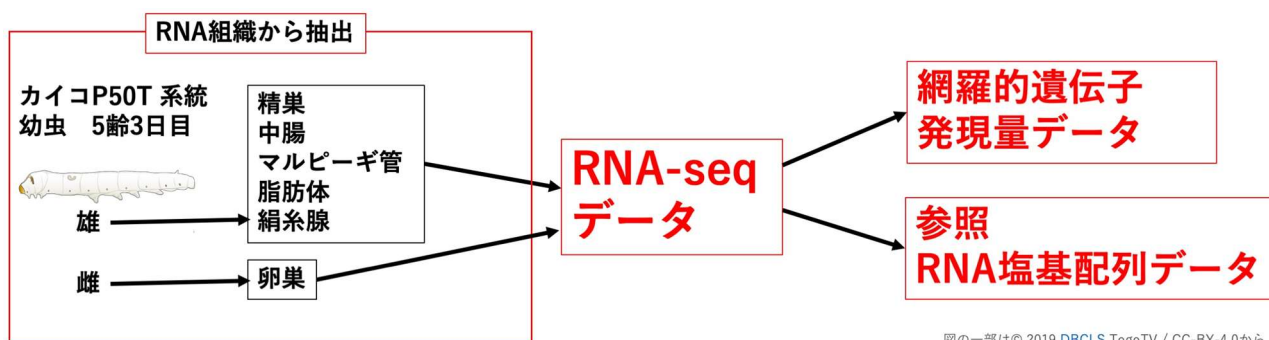


図1 発現データの取得

絹糸の生成に重要な時期である5齢3日目の幼虫の主要な組織からRNAを抽出し、それぞれの組織における遺伝子発現データ、を取得にさらに参照RNA塩基配列を取得しました。

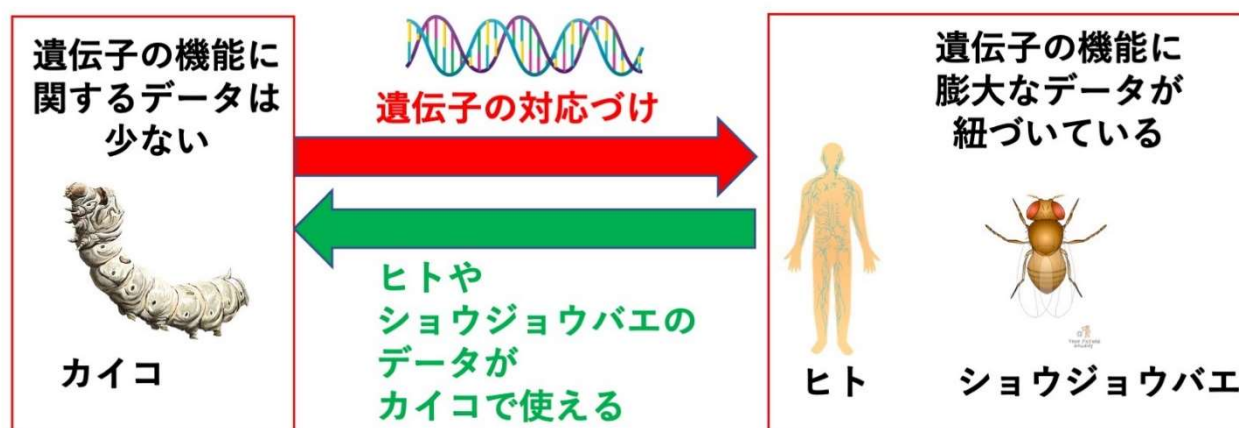
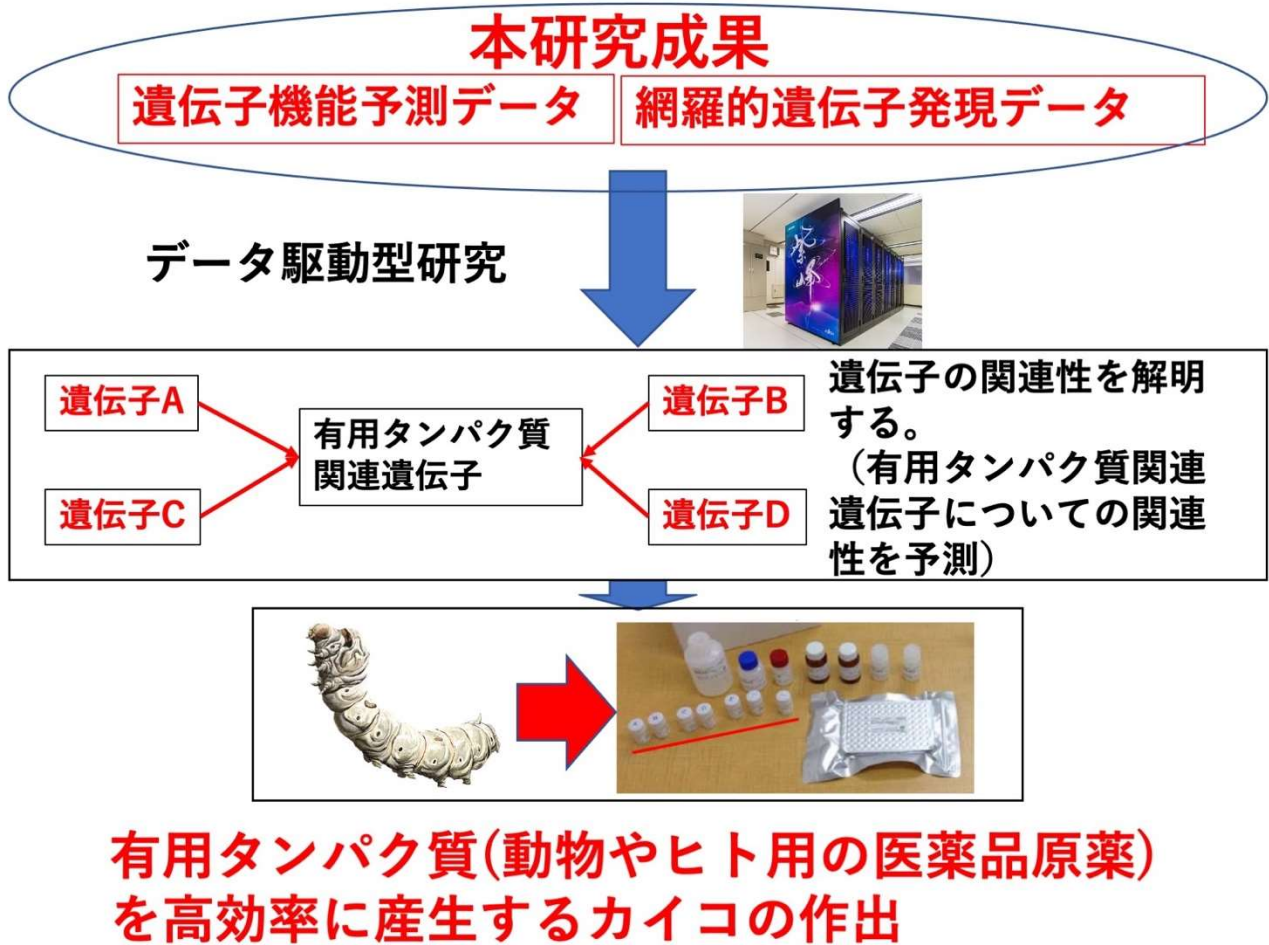


図2 ヒトやショウジョウバエとの比較

カイコの遺伝子に関するデータが少ないため、カイコの遺伝子をヒトやショウジョウバエの遺伝子と対応づけすることで、ヒトやショウジョウバエの遺伝子に紐づいている膨大なデータをカイコの遺伝子に利用することを可能にしました。



図の一部は© 2019 DBCLS TogoTV / CC-BY-4.0から

図3 研究成果の活用

本研究成果である網羅的発現量データを用いてデータ駆動型研究を実施します。有用タンパク質を高効率に産生するカイコを作出することで、カイコによる有用タンパク質産生技術の普及を目指します。