

オーム科学におけるデータベース構築とマイニング技術

金谷 重彦^{†1}

奈良先端科学技術大学院大学・情報科学研究科・計算システムズ生物学研究室

ヒトは動物であるので毎日食事を摂ることによって恒常性を維持する。恒常性を維持するためのシステムとして、料理を例にオミックス科学を説明し、オミックス科学におけるデータベース、特にメタボロームを中心としたデータベースシステム KNApSAcK Family DB およびマイニング技術について紹介する。

Bio-database and Data-mining in Omics: KNApSAcK Family DB

Shigehiko Kanaya^{†1}

To understand interaction between food ingredients and human in a systems-level, we need to accumulate information on biological entities derived from a global analysis of gene regulation, metabolic pathway flux and homeostatic control mechanisms, and develop data-mining techniques in handling the huge data set concerning to them. KNApSAcK Family database (DB) system has been developed by taking into consideration -omics information on the interaction of human and cuisine including food ingredients such as vegetables, fungi, meats and fishes. Cuisines are complex systems consisting of food ingredients and directly affect the human health. We introduce three DBs, processed food DB consists of ingredients of retortable pouched pack of curry (called FoodProcessor; <http://kanaya.naist.jp/FoodProcessor/top.jsp/>), an assessment system of how a cuisine, i.e., a combination of food ingredients affects human health (DietDish, <http://kanaya.naist.jp/DietDish/top.jsp/>) and a retrieval system for finding foods useful in noncommunicable disease prevention (DietNavi, <http://kanaya.naist.jp/DietNavi/top.jsp/>) which are part of KNApSAcK family database (http://kanaya.naist.jp/KNApSAcK_Family/).

ヒトは動物であるので毎日食事を摂ることによって恒常性を維持する。恒常性を維持するためのシステムとして、料理を例に、オミックス科学、特にメタボロームを中心としたこれらのデータベースシステム KNApSAcK Family DB、およびそのマイニング技術について紹介する。

食事は、ヒトの恒常性を維持において最も重要な因子のひとつである。そこで、料理をゲノム研究で考えよう。料理は、複数の食材から構築される複雑系である。「料理を科学する」にはひとつには素材の機能性を考慮した栄養学アプローチおよび薬膳に代表される医薬学アプローチなどが考えられるが、オミックス科学はこの医薬学アプローチに代表される分子メカニズムによりヒトの恒常性を悉皆的に理解するためのひとつの方法ともなり得る。まずはじめに、カレーライス为例に、オミックス科学を説明しよう。

1990 年後半、生物を遺伝子の集合体として理解することを目的に、全ゲノム配列の決定が進められた。2003 年に、ヒトの全ゲノム配列が決定され、ヒトを遺伝子の集合として記述することが可能となった。その直後に、シドニー・ブレナー博士が日本で講演をしたときに、面白いジョークを言った¹⁾。

「DNA は何の訳だと思う？」

「Don't (K)Now Anything.

何もわからなくてこと。

DNA だけですべて説明できるほど、

人間は単純じゃない。だからおもしろい。」

つまり、ゲノム配列のみではヒトを理解することはできず、組織・細胞ごとに動的に発現する全ての遺伝子をもとに動的に構築される生体分子を要素に、ヒトをシステムとして理解することを目標とした研究が、ゲノミクスを基盤とした研究として展開されている。そこで、登場したのがトランスクリプトーム、プロテオーム、メタボロームである。ここで、オームとは全体と言う意味であり、その学問体系をオーム科学もしくはオミックスと呼ぶ。

カレーライスを例に用いてオミックスを整理すると図 1 のようになる。ヒトの生理活性を右側に、薬用・食用の知識を左側に配置した。ヒトは動物であり、食糧を摂取しないと生きていくことができない。すなわち、自分自身の恒常性を維持するために、有用生物を食物として摂取しなければならない。いま、カレーライスを摂取したとしよう。カレーライスには伝統知識が詰め込まれており、これが薬用・食用知識ベースであり、種々の食材（生物）の配合方法が格納される。図 1 では、クミン、コリアンダー、唐辛子などの情報がこの配合にあたる。それぞれの食材生物には機能性代謝物が含まれる。例えば、クミンについては、ク

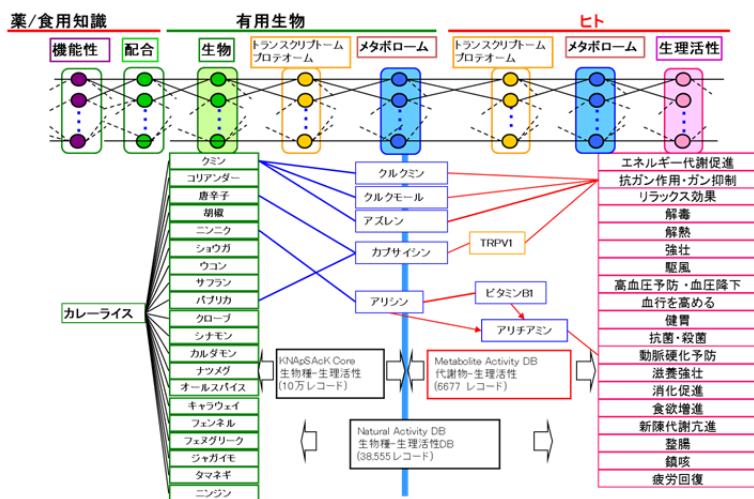


図1 カレーライスを例としたオーム科学の説明図

ルクミン、クルクモールなどである。食品としての安全性を確保し、これらの機能性代謝物の収量を上げるために、作物のゲノム情報をもとにトランスクリプトーム（細胞全体の遺伝子発現）、プロテオーム（細胞全体のタンパク質）、さらにはメタボローム（細胞全体の代謝物）を最適化することが考えられる。このようにゲノムからメタボロームに至るすべての生体要素により細胞・組織全体を理解する研究分野がマルチ・オミックス(オミックス科学)と呼ばれるアプローチである。

カレーライスに含まれる食材を摂取すると、ヒトへの効能としては、エネルギー代謝促進、抗がん作用などがある(図1右側)。この効能は、おおまかには、食材としての生物から摂取した二次代謝物とヒトのタンパク質の相互作用により、ヒトの代謝が動的に最適化されることにより恒常性が維持される。いま、トウガラシの成分であるカプサイシンはタンパク質 TRPV1 を刺激することが知られており、その結果、ヒトの体内で種々のトランスクリプトーム、プロテオーム、メタボロームが動的に変動し抗がん作用、がん抑制としての機能などを発現することになる。このように、薬用・植物知識、有用生物のオミックス情報、ヒトのオミックス情報を統合的に扱うことにより、食品の機能性を分子レベルで理解することが可能となる。そこで、図1に示した薬用・植物知識ベース(機能性、配合)有用生物オミックス情報(生物、ゲノム、トランスクリプトーム、プロテオーム、メタボローム)、ヒトオミックス情報(ゲノム、トランスクリプトーム、プロテオーム、メタボローム)さらにヒト生理活性の大量情報を整理するデータベース(DB)構築ならびに、これらの大量情報を悉皆的に理解することが必要となる^{2,3)}。そこで筆者らの研究室では、メタボロームを中心にデータベースの開発を進めている。

図2にKNAPSAcK Family DBの概要を示す⁴⁻¹¹⁾。また、図3にメインウィンドウを示す。まずはじめに、人類の歴史により培われた食/薬用植物における知識を活用し、生薬

においては日本における漢方薬(KAMPO DB、278 薬用生物、336 処方)、インドネシアにおけるジャム(JAMU DB、1133 薬用植物、5310 処方)といったヒトの体調により合わせて摂取すべき食/薬用植物の情報を体系化しDBに整理した。また、食用に使用されている生物を食用 DB(Lunch Box、709 生物)、さらにはハーブとして利用されている生物をハーブ DB(Tea Pot)として整理し公開した。また、国ごとの食/薬用植物の利用を把握する目的で、世界の薬用植物データベースを構築した(WorldMap DB、222 地域、50,000 地域-薬/薬用植物の関係)。さらに生薬を含む生物において生合成される代謝物を科学文献より網羅し、生物種-代謝物関係 DB(KNAPSAcK Core、

101,500 対の生物種-代謝物の関係、5 万代謝物種、2 万生物種)を構築した。例えば、タマネギは食欲増進に効果があるなど、その生物のヒトに対する活性および抗菌作用のようにその生物が他の生物に対してどのような活性があるかが報告されている。このような、生物が他の生物へ与える活性を生物活性 DB(Biological Activity DB、2,418 種の生物活性における 3 万対の植物とその活性の関係)として整理・公開した。同様に、個々の代謝物が生物へ与える影響を代謝物活性 DB(Metabolite Activity DB)として整理した。また、薬用・植物における有効成分の生合成過程を把握する目的で、植物-酵素-酵素反応からなるデータベース(Motorcycle DB)の構築を進めている。

最近、サイエンス誌に生活習慣病を如何に防ぐかという特集が組まれた¹²⁾。その中で、2050 年までに生活習慣病と関連する人は全人口の約 1/7 と推定されており、生活習慣

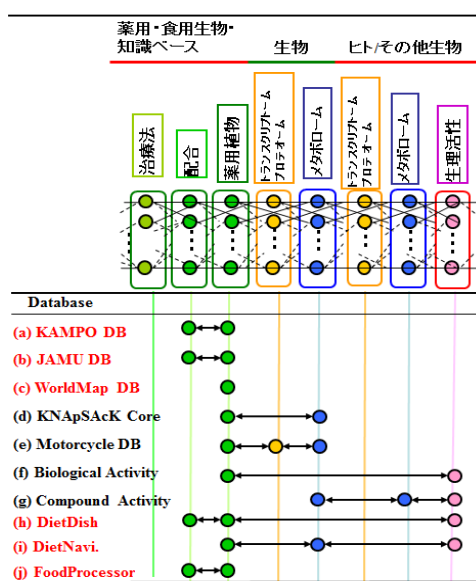


図2 KNAPSAcK Family DBにおける属性観の関係図

病の予防は食事をいかにコーディネートするかが課題である。現在、レトルト・カレーを含む加工食品の普及を考えると、野菜や魚などの生鮮食品のみならず加工食品を含めて食材と生活習慣病の関係を整理すべきである。

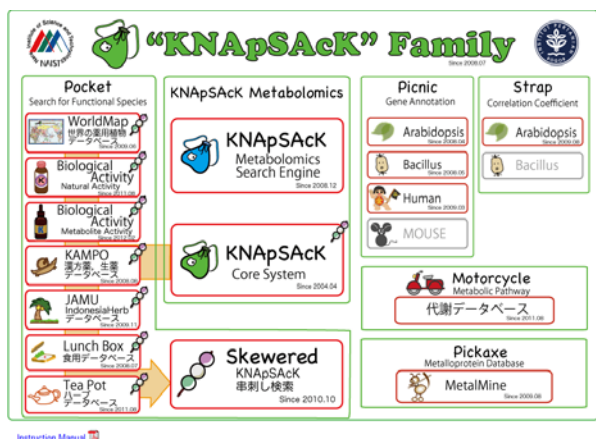


図3 KNApSAcK Family DB のメインウィンドウ

食材には食べ合わせというものがああり、栄養学、薬膳、医薬学で研究がなされ提案されている。例えば、キュウリはビタミンCを分解する酵素を有するため、ビタミンCが豊富に含まれる食材と食べ合わせが悪いと言われている。そこで、料理の食べ合わせを評価するためのシステム DietDish DB を構築した。2012年9月4日現在、414 食材について、3708 組の食材食べ合わせデータが格納されており、料理の中に含まれる二つの食材に食べ合わせ効果が報告されていれば2点、三つの食材について食べ合わせ効果が報告されていれば3点、逆に、二つの食材に食べ合わせに負の効果が報告されていれば-2点として、料理の食材群についての全ての食材の組み合わせをもとに合計点を求めるシステムである。そこで、DietNavi の開発を進めている。このシステムは、肥満、高血圧、高脂血症などの生活習慣病を防ぐにはどのような食材および食材の成分が有効であるかを提供するためのシステムである。このようにオミックス科学で必要とされるデータは、ゲノムからはじまり、最終的にはヒトの健康という社会への接点までを研究対象とする。そこでは、いわゆる生物学・栄養学・薬学・医学情報を体系化するための情報科学がなくしては研究が達成できない。講演に際しては、オミックス科学におけるデータマイニング法についても紹介したい。

謝辞 本研究は JST-NBDC、科学研究費・新領域・バイオマシナリープロジェクトなどの支援のもとに進めている。ここに感謝する。

参考文献

1)本村有紀子、理系思考 分からないから面白い、毎日新聞社 (2007)

2) T. Okada, F. M. Afendi, M. Altaf-UI-Amin, H. Takahashi, K. Nakamura, and S. Kanaya, *Curr Comput Aided Drug Des.*, 6, 179-196, (2010)

3) F. M. Afendi, T. Katsuragi, A.Kato, N. Nishihara, K. Nakamura, Y. Nakamura, K. Tanaka, A. H. Morita, M. Altaf-UI-Amin1, H. Takahashi, and S. Kanaya, *Current Pharmacogenomics and Personalized Med.*, 10, 111-124 (2012)

4)Y. Shinbo, Y. Nakamura, M. Altaf-UI-Amin, H. Asahi, K. Kurokawa, M. Arit, K. Saito, D. Ohta, D. Shibata amd S. Kanaya, *Biotechnology in agriculture and forestry*, 57, 166-181 (2006)

5)F. M. Afendi, T. Okada, M. Ymazaki, A. H. Morita, Y. Nakamura, K. Nakamura, S. Ikeda, H. Takahashi, M. Altaf-UI-Amin, L. K. Darusman, K. Saito and S. Kanaya, *Plant Cell Physiol.*, 53, e1-e12, (2012)

6) 金谷重彦、西岡孝明、有田正規、櫻井望、細胞工学、1, 101-104 (2012)

7) 金谷 重彦、平井 晶、旭 弘子、高橋弘喜、中村建介、Md. Altaf-UI-Amin, 二瓶 義人、池田 奨、尾島 雄也、西岡孝明、実験医学, 29, 2460-2470 (2012)

8) 中村由紀子、森田(平井)晶、西岡孝明、金谷重彦、バイオサイエンスとインダストリー、70, 267-272 (2012)

9)池田俊、桂樹哲雄、小野直亮、中谷淳至、中村由紀子。森田晶、金谷重彦、生物工学、90、(印刷中)(2012)

10) 森田(平井)晶、中村由紀子、桂樹徹、金谷重彦、*Medical Herb*, 20, 20-23 (2012)

11) 桂樹哲雄、小野直亮、森田(平井)晶、中村由紀子、Md.Altaf-UI-Amin、金谷重彦、*FFI ジャーナル* (印刷中) (2013)

12) C. Ash, P. Kiberstis, E. Marshall, J. Travis, *Sci.*,337, 1467, (2012)