

独立行政法人科学技術振興機構
バイオインフォマティクス推進センター事業

創造的な生物・情報知識融合型の研究開発

平成19年度採択研究開発課題

(追跡調査報告書)

研究開発課題：

酵素反応分類に基づく酵素反応予測システムの開発
(平成19年10月～平成22年9月)

代表研究者氏名：長野 希美

(産業技術総合研究所 生命情報工学研究センター 主任研究員)

目次

本編

1.	研究開発による成果.....	1
1.1.	研究開発課題の目標及び新規性・独創性.....	1
1.2.	研究開発開始時の課題を取り巻く研究・技術水準及び分野における課題の位置づけ.....	1
1.3.	研究開発終了時の成果概要.....	1
1.4.	研究開発の達成度.....	2
2.	研究開発による成果の活用状況や発展状況.....	3
2.1.	構築されたデータベース・ソフトウェア等の活用状況.....	3
2.1.1.	データベース・ソフトウェア等の継続状況や発展状況.....	4
2.1.2.	第三者によるデータベース・ソフトウェア等の活用事例.....	4
2.1.3.	データベース・ソフトウェア等へのアクセス数／ダウンロード数、visit 数... ..	4
2.2.	課題終了後の研究開発成果の継続状況や発展状況.....	5
2.2.1.	BIRD 終了から現在まで、BIRD で取り組んだ課題に関連した研究開発成果の 継続状況や発展状況（国内外の研究者との共同研究へ発展した等）.....	5
2.2.2.	BIRD 終了後に発表された論文リスト.....	5
2.2.3.	研究開発成果の現在の国際的な評価・位置づけ.....	6
2.3.	現在の科学技術研究における研究開発成果の活用状況、発展状況のまとめ.....	6
3.	研究開発による成果、の科学技術的、社会的及び経済的な効果.....	6
3.1.	バイオインフォマティクス分野及びライフサイエンス分野の研究への貢献.....	6
3.1.1.	バイオインフォマティクス分野の研究への貢献.....	6
3.1.2.	ライフサイエンス分野（バイオインフォマティクス以外）や その他、科学技術分野の研究への貢献.....	6
3.2.	人材育成の面から参加研究者の活動状況.....	7
3.2.1.	研究開発に参加した研究者のキャリアアップ.....	7
3.3.	社会的及び経済的な効果.....	7
3.3.1.	研究開発成果が大学や公的研究機関、企業等で、応用に向けて 継承または発展した例.....	7
3.3.2.	その他、研究開発成果が社会的、経済的な効果・効用につながる兆し、可能性.....	7

資料編

1.	論文リスト.....	8
2.	主要論文の被引用回数.....	10
3.	学会招待講演・基調講演.....	12
4.	新聞発表等.....	12

5.	特許出願・成立	12
6.	学会賞等の受賞	12
7.	グラントの獲得実績.....	12
8.	書籍等の執筆実績	12
9.	総説の執筆実績	13
10.	参加研究者の活動状況.....	13

本調査報告書は平成 26 年 4 月に作成
表紙の代表研究者所属は平成 25 年 10 月時点

1. 研究開発による成果

1.1. 研究開発課題の目標及び新規性・独創性

従来の酵素分類EC番号は、主に基質・産物の化学構造や補酵素などに基づいた分類であり、触媒機構において重要な蛋白質の配列情報、立体構造は考慮されていない。本課題では、EC番号に替わる、より精密で階層的な酵素反応分類を網羅的に行うことを目標とした。これは詳細な酵素触媒機構、立体構造、リガンドの反応部位に基づいた酵素反応クラスの新しい階層分類である。また、分類を促進するために、独創性の高い機械学習の最新技術を駆使した自動酵素反応予測システムの開発、また、この予測システムを用いた機能未知酵素の酵素反応クラスの予測とその結果の分類への反映、最終的には「酵素反応予測システム」の一般公開を目指した。

本課題開始前に開発した EzCatDB (Enzyme Catalytic Mechanisms : 酵素触媒機構データベース) は、当時から一般公開されており、月間 3~4 万件程度の外部アクセスがあった。このデータベースをさらに使いやすく整備し、新しい予測解析ツールを追加することにより、ユーザーが利用しやすく独創性の高いシステムに改良した。

1.2. 研究開発開始時の課題を取り巻く研究・技術水準及び分野における課題の位置づけ

酵素は、創薬など産業応用上重要な蛋白質であり、5千種類程あるとされ、多種多様な生体反応を担う。従来の酵素の分類であるEC番号は、主に基質・産物の化学構造や補酵素などに基づき分類しており、酵素のID番号として世界的に普及している。しかし、EC番号では、途中過程の触媒反応において重要な蛋白質の配列、立体構造が考慮されていなかった。こうした問題点を抱えながらも、多くの場合、蛋白質機能予測はEC番号を基に、配列解析により行われていた。

こうした問題点を踏まえ、本研究グループではEzCatDBを開発し、EC番号とは異なる、独自の酵素反応階層分類を行い、進化系統と反応の関連などの研究を進めた。

一方、構造ゲノミクスで出されている膨大な蛋白質立体構造のデータに対する機能予測法の開発も待たれており、そのためにも新しい酵素反応の分類法と予測法の開発が期待されていた。

1.3. 研究開発終了時の成果概要

【相同反応検索システム】

EzCatDBに登録されている活性部位情報を用いて、配列データを基に、相同反応データ検索システムとしてEzCat-BLASTという新しいツールを開発した。これはEzCatDBデータベース中の配列と活性部位の一致度をウェブ上に表示するシステムであり、活性部位の保存度を表示するものである。保存度が1になった場合、クエリ配列は、ヒットした配列に対して完全に保存された活性部位残基を持つ。

【類似反応検索システム】

相同反応データ検索と異なり、類似反応の探索は配列比較では難しいため、活性部位の局所立体構造の比較が必要となる。局所構造の比較によって活性部位を予測する際、従来はRMSDと呼

ばれる原子間距離を単純に平均した尺度が用いられていた。類似反応を担う酵素の予測研究では、ここに機械学習技術を導入して高精度に活性部位予測を試みた。局所構造比較による活性部位予測を行う際、シングル・テンプレート予測やマルチプル・テンプレート予測と呼ばれる2種類の解析があるが、これら2種類の解析のための算法を開発した。

RMSDは平均二乗残差に平方根をとったものであるが、原子に重みづけせずに距離の平均をとっている(UMD)。本研究開発では重み付きの距離(WMD)を提案している。もう一つのアプローチとして、DALIスコアを線形結合した値を類似度とする尺度も開発した(DSDS)。WMDの最適パラメータを自動計算するために計量学習アルゴリズム、DSDSの最適パラメータを学習するために正則化ロス最小化アルゴリズムを開発した。また、DSDSの機械学習の効果を確認するために、機械学習を用いないDALIスコアの平均(MDS)とも比較した。

実際の立体構造を使った予測実験の結果は、シングル・テンプレート予測、マルチプル・テンプレート予測のどちらの場合も、本研究で使用されたアルゴリズムが従来法を有意に上回る性能の向上を得た。この予測アルゴリズムを基に、プロトタイプシステムを開発した。

以上述べたように、類似反応予測法に関しては、本研究開発終了時点で、アルゴリズムの構築、実験、プロトタイプのシステムの構築まで終え、論文発表も行い、3報が受理されている。この類似反応検索をWEB上で行うシステムであるEzMetActはこの研究開発終了時点ではまだ開発中であり、EzCat-BLASTと同時の公開を目指していた。

【EzCatDB overviewサイト】

EzCatDB整備の一環として新規のサイト、「overviewサイト」を作成した。これは①酵素反応分類の全体を示すリスト、②触媒反応と対応する酵素のリスト、③触媒機構の種類と対応する酵素のリストをまとめたものであり、本データベースにおける酵素分類の全体像をわかりやすく示すものである。

1.4. 研究開発の達成度

当初の計画のメインテーマは「相同反応」及び「類似反応」の酵素反応予測法の開発を行うことであった。相同反応予測法に関しては、EzCat-BLASTの開発まで進んでいる。また類似反応予測法に関しては、アルゴリズムの構築、実験、プロトタイプのシステムの構築を行い、論文発表も行なった。

一方、本開発を行うために計量学習という新しい枠組みを取り入れることにより、局所構造比較・予測法における偽陽性のデータの数を大きく減らし、従来は非常に難しかった局所構造比較法による類似した活性部位を探索する事が可能になった。また確率モデルにより偏差距離のスコア変換を取り入れることにより、多くの異なる局所構造を統一的な基準で比較する事ができるようになり、マルチプル・テンプレート解析が可能になった。計量学習と確率モデルによるスコア変換の組合せは、計量学習しない場合の確率モデルによるスコア変換の結果よりも、良好な結果を得た。

また、酵素の生物種分布の解析は当初の研究計画にはなかったが、相同酵素の解析システムの

開発とともに、得られた相同酵素の生物種分布を解析し、生物特異的な酵素や生物種普遍的に存在するユビキタス酵素を求めることができるようになり、新しい研究展開への可能性を示すことができた。

2. 研究開発による成果の活用状況や発展状況

2.1. 構築されたデータベース・ソフトウェア等の活用状況

EzCat-BLAST

本プログラムは代表研究者により作成された酵素データベース、EzCatDBの配列のBLAST検索を行うためのものであり、EzCatDBに登録されている配列のみならず、Swiss-Prot/UniProtに登録されている酵素などの配列について活性部位の配列を考慮した検索も行い、機能部位の保存度などを表示することができるものである。本研究開発終了時点においては一般公開する予定となっていたが、配列を検索するシステムは導入されていない。より高感度な配列検索システムに改良しているところであり、2014年夏までに一般公開される予定である。

EzCat-BLASTとEzCatDBが、微生物の二次代謝物の生合成メカニズムの解明などに応用されている事例として、カビにおける二次代謝物の遺伝子クラスターを解析した一部の結果は、Umemura *et al.*, PLoS One. Vol 8(12), 2013, e84028で発表されている。また、糸状菌（カビ）では、従来、ペプチド系の二次代謝物は、非リボソームペプチド合成系（NRPS）により合成されているとされていたが、EzCat-BLASTなどの活用により、NRPSを使わないリボソームペプチドも存在することを発見するに至った（特願2014-044277; Umemura *et al.*, Fungal Genomics & Biologyに投稿中）。

EzMetAct

PDB IDを入力するかPDBフォーマットのデータをアップロードすると、計量学習により得られた距離スコアと、PINTSスコアと共に、該当するテンプレートに対応するアミノ酸残基で示す部位のリストと立体構造を表示するものである。本研究でEzMetActのプロトタイプの開発は成功したが、本機能の公開システムは導入されていない。実用に耐えるだけテンプレートの数を増やしてから、公開される予定である。

EzCatDB overviewサイト

本研究開発においてEzCatDBに加えられた新規サイトであり、EzCatDB中で酵素反応分類がされているエントリのリスト、(2)触媒反応を分類し該当する酵素例を示すリスト、(3)加水分解反応、転移反応の触媒機構の種類と該当する酵素のリストの3つのページが含まれる。EzCatDBの分類とそれぞれの酵素反応についての情報の全体が分かりやすい形で示されており、また外部とのリンクもとりやすいようになっている。EzCatDBのトップページに本サイトへのボタンが分かりやすく示されており、本DBに初めてアクセスした場合にも役に立つページを提供している。

2.1.1. データベース・ソフトウェア等の継続状況や発展状況

	ツール名	URL	公開状況	備考
(1)	EzCat-BLAST	—	未公開	2013 年 末 時 点 に お い て EzCatDB は更新されているものの、配列を検索するシステムは導入されていない。BIRD 終了後もその研究に関する報告はない。ただし、より高感度な配列検索システムを構築しているため、2014 年の夏までに一般公開される予定である。
(2)	EzMetAct	—	未公開	2013 年 末 時 点 に お い て EzCatDB は更新されているものの、本プログラムを活用したシステムは導入されていない。本プログラムは開発者により継続して使用されているが、一般公開はされていない。
(3)	EzCatDB overview	http://mbs.cbrc.jp/EzCatDB/overview/	公開中（無料）	

2.1.2. 第三者によるデータベース・ソフトウェア等の活用事例

特になし。

2.1.3. データベース・ソフトウェア等へのアクセス数／ダウンロード数、visit 数

	ツール名	アクセス数、ダウンロード数、visit 数	集計期間	備考
(1)	EzCat-BLAST	未公開		
(2)	EzMetAct	未公開		
(3)	EzCatDB overview	2,360	2012 年 1 月～2013 年 12 月	研究開発終了報告書

2.2. 課題終了後の研究開発成果の継続状況や発展状況

2.2.1. BIRD 終了から現在まで、BIRD で取り組んだ課題に関連した研究開発成果の継続状況や発展状況（国内外の研究者との共同研究へ発展した等）

EzMetActは一般公開されていないが、本研究のプログラムに関する研究は代表研究者の報告として発表されており現在も継続している。

2.2.2. BIRD 終了後に発表された論文リスト

T. Kato, K. Suwa, N. Nagano: Parametric Templates: A New Enzyme Active-Site Prediction Algorithm. In conjunction with the 10th IEEE International Conference on Data Mining (ICDM 2010) Biological Data Mining and its Applications in Healthcare, 711-718 (2010.12.13)

T. Kato, N. Nagano: Discriminative Structural Approaches for Enzyme Active-Site Prediction. BMC Bioinformatics, presented at The Ninth Asia Pacific Bioinformatics Conference (APBC2011 Incheon, Korea. 11-14 January 2011), BMC Bioinformatics 2011, 12(Suppl 1):S49

Motono, C., Nakata, J., Koike, R., Shimizu, K., Shiota, M., Amemiya, T., Tomii, K., Nagano, N., Sakaya, N., Misoo, K., Sato, M., Kidera, A., Hiroaki, H., Shirai, T., Kinoshita, K., Noguchi, T., Ota, M.: SAHG, a comprehensive database of predicted structures of all human proteins, Nucleic Acids Research, Volume 39, Issue SUPPL. 1, January 2011, Pages D487-D493

Nagao C, Nagano N, Mizuguchi K: Prediction of detailed enzyme functions and identification of specificity determining residues by random forests. PLoS One. Volume 9, Issue 1, January 2014, Page e84623

Umemura M, Koike H, Nagano N, Ishii T, Kawano J, Yamane N, Kozono I, Horimoto K, Shin-ya K, Asai K, Yu J, Bennett JW, Machida M: MIDDAS-M: motif-independent de novo detection of secondary metabolite gene clusters through the integration of genome sequencing and transcriptome data. PLoS One. Volume 8, Issue 12, December 2013, Page e84028

Raissa Relator, Tsuyoshi Kato, Richard Lemence: Improved protein-ligand prediction using kernel weighted canonical correlation analysis, IPSJ Transaction on Bioinformatics, Volume 6, June 2013, Pages 18-28

Reiji Teramoto and Tsuyoshi Kato: Transfer learning for cytochrome P450 isozyme selectivity prediction, Journal of Bioinformatics and Computational Biology, Volume 9, Issue 4, 2011, Pages 521-540

Jaak Simm, Masashi Sugiyama, and Tsuyoshi Kato, Computationally efficient multi-task learning with least-squares probabilistic classifiers. IPSJ Transactions on Computer Vision and Applications, Volume 3, 2011, Pages 1-8

2.2.3. 研究開発成果の現在の国際的な評価・位置づけ

研究開発終了報告書において主要論文とされた5報のうち、被引用情報が取得できた3報についてみると、いずれも海外からも引用されており、本課題の成果が国内に留まらないことを示していると言えるが、2012年12月までのデータでは被引用件数自体はまだ少なく、被引用情報だけで国際的な評価・位置づけを判断することはできない。

2.3. 現在の科学技術研究における研究開発成果の活用状況、発展状況のまとめ

本研究開発で作成したプログラムを導入する対象となる酵素反応データベースであるEzCatDBはユニークなデータベースであり、開発されたプログラムの公開と外部による利用が期待された。しかし、本研究で中心的開発課題であるEzCat-BLASTとEzMetActは公開されていない（2013年末時点）ため、その活用は論文として発表された情報に止まっている。ただし、これまで、サーバー、OSなどのシステム担当者の変更により、メンテナンス及び開発が遅れていたが、EzCat-BLASTよりも高感度なアルゴリズムを採用した配列検索システムが開発されたので、EzCatDBの新システムと共に、2014年夏頃までに、一般公開される予定である。（夏頃までとしているのは、NARのデータベースissueに投稿予定であるため。）

3. 研究開発による成果、の科学技術的、社会的及び経済的な効果

3.1. バイオインフォマティクス分野及びライフサイエンス分野の研究への貢献

3.1.1. バイオインフォマティクス分野の研究への貢献

本課題では、普遍性の高い類似反応の探索理論を構築し、いくつかの例題を解くことに成功している。またこの理論に基づき開発されたシステムをタンパク質立体構造データの機能解析などに応用することにより、生体内におけるタンパク質の役割などの網羅的理解に役立つ可能性がある。

3.1.2. ライフサイエンス分野（バイオインフォマティクス以外）やその他、科学技術分野の研究への貢献

従来問題の多かった酵素の分類を見直し、新たな観点から分類したデータベースを整備し、その予測システムを開発したことは酵素を取り扱う研究者にとって有意義である。精度の良い酵素反応予測法は、構造ゲノミクス・プロジェクトの成果からライフサイエンスにとって有用な情報を引き出す上で不可欠であり、本課題の成果発表が及ぼす影響は小さくないと思われる。しかし、予測システム自体は現在まで公開されておらず、その意義は現時点では不明である。

酵素反応予測自体が非常に困難な課題であり、さらに精度の高い予測法を開発し、公開することにより評価を得る事は極めて重要である。一方、今回開発したプログラムの対象となるデータベースであるEzCatDBはユニークなデータベースであり、その継続とともにゲノム情報との連携

が求められており、さらなる反応機構データの拡張と予測機能の充実が期待される。

EzCatDBと配列検索システムEzCat-BLASTが、微生物の二次代謝物の生合成メカニズムの解明などに応用されている。微生物、植物などの二次代謝生合成機構は未知の部分が多いため、こうした酵素反応のデータベースや配列検索システムが重要な役割を果たすと考えられる。

3.2. 人材育成の面から参加研究者の活動状況

3.2.1. 研究開発に参加した研究者のキャリアアップ

本研究開発課題に参加した者で、その後新たなポジションを得た者は3名いた。

3.3. 社会的及び経済的な効果

3.3.1. 研究開発成果が大学や公的研究機関、企業等で、応用に向けて継承または発展した例

現時点でそのような情報は得られていない。

3.3.2. その他、研究開発成果が社会的、経済的な効果・効用につながる兆し、可能性

EzCat-BLAST が公開され、酵素活性部位の立体構造と触媒機能の相関関係を求めることが可能になれば、世界各国で実施されている構造ゲノミクス・プロジェクトにより解かれている膨大な蛋白質の立体構造を用いて、機能未知の蛋白質の機能予測を行うことができる。

また、相同反応、類似反応の網羅的解析から、酵素反応の選択性などを推測する研究へと発展させることも可能になる。更に、酵素反応の選択性の解析から酵素に作用するリガンド化合物の生体内での副作用のメカニズム解明などへの応用も期待される。

資料編

1. 論文リスト

1. T. Kato, N. Nagano: Metric Learning for Enzyme Active-Site Search. *Bioinformatics*, 26, 2698-2704 (2010).
2. *T. Kato, K. Suwa, N. Nagano: Parametric Templates: A New Enzyme Active-Site Prediction Algorithm. In conjunction with the 10th IEEE International Conference on Data Mining (ICDM 2010) Biological Data Mining and its Applications in Healthcare, 711-718 (2010.12.13)
3. *T. Kato, N. Nagano: Discriminative Structural Approaches for Enzyme Active-Site Prediction. *BMC Bioinformatics*, presented at The Ninth Asia Pacific Bioinformatics Conference (APBC2011 Incheon, Korea. 11-14 January 2011), *BMC Bioinformatics* 2011, 12(Suppl 1):S49
4. T. Kato, H. Kashima, M. Sugiyama, K. Asai: Multi-task learning via conic programming, In J. C. Platt, D. Koller, Y. Singer, and S. Roweis (Eds.), *Advances in Neural Information Processing Systems*, 20, 737-744, Cambridge, MA, MIT Press, 2008, presented at Twenty-First Annual Conference on Neural Information Processing Systems (NIPS2007)
5. *T. Kato, H. Kashima, M. Sugiyama, K. Asai: Conic Programming for Multi-Task Learning, *IEEE Transactions on Knowledge and Data Engineering*, 22(7), 957-968 (2010).
6. C. Nagao, N. Nagano, K. Mizuguchi: Relationships between functional subclasses and information contained in active-site and ligand-binding residues in diverse superfamilies. *PROTEINS: Structure, Function, and Bioinformatics*, 78, 2369-2384 (2010).
7. *加藤 毅, 諏訪 和大, 長野 希美: 計量学習による酵素活性部位予測, 電子情報通信学会技術研究報告 110(83), 103-110
8. T. Kato, H. Kashima, M. Sugiyama, and K. Asai: Robust Label Propagation on Multiple Networks, *IEEE Transactions on Neural Networks*, Vol.20, No.1, pp.35-44.2009
9. Motono, C., Nakata, J., Koike, R., Shimizu, K., Shiota, M., Amemiya, T., Tomii, K., Nagano, N., Sakaya, N., Misoo, K., Sato, M., Kidera, A., Hiroaki, H., Shirai, T., Kinoshita, K., Noguchi, T., Ota, M.: SAHG, a comprehensive database of predicted structures of all human proteins, *Nucleic Acids Research*, Volume 39, Issue SUPPL. 1, January 2011, Pages D487-D493
10. Nagao C, Nagano N, Mizuguchi K: Prediction of detailed enzyme functions and identification of specificity determining residues by random forests. *PLoS One*. Volume 9, Issue 1, January 2014, Page e84623
11. Umemura M, Koike H, Nagano N, Ishii T, Kawano J, Yamane N, Kozono I, Horimoto K, Shin-ya K, Asai K, Yu J, Bennett JW, Machida M: MIDDAS-M: motif-independent de

novo detection of secondary metabolite gene clusters through the integration of genome sequencing and transcriptome data. PLoS One. Volume 8, Issue 12, December 2013, Page e84028

12. Raissa Relator, Tsuyoshi Kato, Richard Lemence: Improved protein-ligand prediction using kernel weighted canonical correlation analysis, IPSJ Transaction on Bioinformatics, Volume 6, June 2013, Pages 18-28
13. Reiji Teramoto and Tsuyoshi Kato: Transfer learning for cytochrome P450 isozyme selectivity prediction, Journal of Bioinformatics and Computational Biology, Volume 9, Issue 4, 2011, Pages 521-540
14. Jaak Simm, Masashi Sugiyama, and Tsuyoshi Kato, Computationally efficient multi-task learning with least-squares probabilistic classifiers. IPSJ Transactions on Computer Vision and Applications, Volume 3, 2011, Pages 1-8

* : 研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの

2. 主要論文の被引用回数

論文	国内外別 件数		分野別件数																						出版年別件数						
	国内 件数	海外 件数	農学	生物学・ 生化学	化学	臨床 医学	コン ピュ ー タ サイ エン ス	経済 学・ 経営 学	工学	環 境・ 生態 学	地球 科学	免疫 学	材料 科学	数学	微生物 学	分子 生物学 ・ 遺伝 学	複合 領域	神経 科学 ・ 行動 科学	薬理 学・ 毒物 学	物理 学	植物 学・ 動物 学	精神 医学 ・ 心理 学	社会 科学 ・ 一般	宇宙 科学	2006 年	2007 年	2008 年	2009 年	2010 年	2011 年	2012 年
1. *T. Kato, N. Nagano: Metric Learning for Enzyme Active-Site Search. Bioinformatics, 26, 2698-2704 (2010).(査読あり)	0	2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	0	0	2
3. *T. Kato, N. Nagano: Discriminative Structural Approaches for Enzyme Active-Site Prediction. BMC Bioinformatics, presented at The Ninth Asia Pacific Bioinformatics Conference (APBC2011 Incheon, Korea. 11-14 January 2011), BMC Bioinformatics 2011, 12(Suppl 1):S49 (査読あり)	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	0	1
5. *T. Kato, H. Kashima, M. Sugiyama, K. Asai: Conic Programming for Multi-Task Learning, IEEE Transactions on Knowledge and Data Engineering, 22(7), 957-968 (2010). (査読あり)	2	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	1	0	2
6. C. Nagao, N. Nagano, K. Mizuguchi Relationships between functional subclasses and information contained in active-site and ligand-binding residues in diverse superfamilies. PROTEINS: Structure, Function, and Bioinformatics, 78, 2369-2384 (2010). (査読あり)	0	2	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	1	1	0
8. Tsuyoshi Kato, Hisashi Kashima, Masashi Sugiyama, and Kiyoshi Asai: Robust Label Propagation on Multiple Networks, IEEE Transactions on Neural Networks, Vol.20, No.1, pp.35-44. (査読あり)	2	5	0	0	0	0	2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	2	0	1	4
9. Motono, C., Nakata, J., Koike, R., Shimizu, K., Shiota, M., Amemiya, T., Tomii, K., Nagano, N., Sakaya, N., Misoo, K., Sato, M., Kidera, A., Hiroaki, H., Shirai, T., Kinoshita, K., Noguchi, T., Ota, M.: SAHG, a comprehensive database of predicted structures of all human proteins, Nucleic Acids Research, Volume 39, Issue SUPPL. 1, January 2011, Pages D487-D493(査読あり)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	0	0
13. Reiji Teramoto and Tsuyoshi Kato: Transfer learning for cytochrome P450 isozyme selectivity prediction, Journal of Bioinformatics and Computational Biology, Volume 9, Issue 4, 2011, Pages 521-540	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	1	0

- ・論文リストについて、トムソン・ロイター社 Web of Science で調査した。調査対象は、2012 年 12 月まで。被引用情報が取得できたものののみ記載した。
- ・論文番号は「資料編 1. 論文リスト」に対応している。
- ・*：研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの。
- ・[国内外別件数] は、被引用文献の国内外別件数。被引用文献の著者の所属国のうち JAPAN が一つでもあれば国内としている。国内外の合計が全被引用文献数となる。
- ・[分野別件数] は、被引用文献の分野別件数。被引用文献が論文の場合のみカウントしている。
- ・[出版年別件数] は、被引用文献の出版年別件数。「—」は、データなしを表す。

論文	Web of Science 国内外別件 数(2012 年 12 月まで)			Google Scholar	
	国内件 数	海外件 数	合計	～2014 年 4 月 24 日	～2012 年
1. *T. Kato, N. Nagano: Metric Learning for Enzyme Active-Site Search. Bioinformatics, 26, 2698-2704 (2010).(査読あり)	0	2	2	4	3

2. * T. Kato, K. Suwa, N. Nagano: Parametric Templates: A New Enzyme Active-Site Prediction Algorithm. In conjunction with the 10th IEEE International Conference on Data Mining (ICDM 2010) Biological Data Mining and its Applications in Healthcare, 711–718 (2010.12.13)	—	—	—	(文献探せず)	(文献探せず)
3. * T. Kato, N. Nagano: Discriminative Structural Approaches for Enzyme Active-Site Prediction. BMC Bioinformatics, presented at The Ninth Asia Pacific Bioinformatics Conference (APBC2011 Incheon, Korea. 11–14 January 2011), BMC Bioinformatics 2011, 12(Suppl 1):S49 (査読あり)	0	1	1	3	3
4. T. Kato, H. Kashima, M. Sugiyama, K. Asai: Multi-task learning via conic programming. In J. C. Platt, D. Koller, Y. Singer, and S. Roweis (Eds.), Advances in Neural Information, Processing Systems, 20, 737–744, Cambridge, MA, MIT Press, 2008, presented at Twenty-First Annual Conference on Neural Information Processing Systems (NIPS2007)	—	—	—	30	24
5. * T. Kato, H. Kashima, M. Sugiyama, K. Asai: Conic Programming for Multi-Task Learning. IEEE Transactions on Knowledge and Data Engineering, 22(7), 957–968 (2010). (査読あり)	2	1	3	15	13
6. C. Nagao, N. Nagano, K. Mizuguchi Relationships between functional subclasses and information contained in active-site and ligand-binding residues in diverse superfamilies. PROTEINS: Structure, Function, and Bioinformatics, 78, 2369–2384 (2010). (査読あり)	0	2	2	4	3
8. Tsuyoshi Kato, Hisashi Kashima, Masashi Sugiyama, and Kiyoshi Asai: Robust Label Propagation on Multiple Networks, IEEE Transactions on Neural Networks, Vol.20, No.1, pp.35–44. (査読あり)	2	5	7	25	16
9. Motono, C., Nakata, J., Koike, R., Shimizu, K, Shiota, M., Amemiya, T., Tomii, K., Nagano, N., Sakaya, N., Misoo, K., Sato, M., Kidera, A., Hiroaki, H., Shirai, T., Kinoshita, K., Noguchi, T., Ota, M.: SAHG, a comprehensive database of predicted structures of all human proteins, Nucleic Acids Research, Volume 39, Issue SUPPL. 1, January 2011, Pages D487–D493(査読あり)	0	0	0	3	0
10. Nagao C, Nagano N, Mizuguchi K: Prediction of detailed enzyme functions and identification of specificity determining residues by random forests. PLoS One. Volume 9, Issue 1, January 2014, Page e84623	—	—	—	(引用元記載なし)	(引用元記載なし)
11. Umemura M, Koike H, Nagano N, Ishii T, Kawano J, Yamane N, Kozono I, Horimoto K, Shin-ya K, Asai K, Yu J, Bennett JW, Machida M: MIDDAS–M: motif-independent de novo detection of secondary metabolite gene clusters through the integration of genome sequencing and transcriptome data. PLoS One. Volume 8, Issue 12, December 2013, Page e84028	—	—	—	2	0
12. Raissa Relator, Tsuyoshi Kato, Richard Lemence: Improved protein–ligand prediction using kernel weighted canonical correlation analysis, IPSJ Transaction on Bioinformatics, Volume 6, June 2013, Pages 18–28	—	—	—	(引用元記載なし)	(引用元記載なし)
13. Reiji Teramoto and Tsuyoshi Kato: Transfer learning for cytochrome P450 isozyme selectivity prediction, Journal of Bioinformatics and Computational Biology, Volume 9, Issue 4, 2011, Pages 521–540	—	—	—	(文献探せず)	(文献探せず)
14. Jaak Simm, Masashi Sugiyama, and Tsuyoshi Kato, Computationally efficient multi-task learning with least-squares probabilistic classifiers. IPSJ Transactions on Computer Vision and Applications, Volume 3, 2011, Pages 1–8	—	—	—	14	8

・ 情報科学・計算機科学分野における研究開発成果では、会議予稿集での発表に引用されることが多いため、Google Scholar での被引用件数を参考情報として添付した。

・ Google での調査方法

対象論文のタイトルを検索し、当該論文を「引用元」とする件数を取得した。

「文献探せず」とあるのはその文献がヒットしなかったもの。

「引用元記載なし」とあるのは、「引用元」の表記がなく引用件数のデータがないもの。

最新の被引用件数と、それから 2013 年以降の被引用件数を引いたものを 2012 年までとした。ただし、年が明確でないものは引かれていないため、2012 年までの被引用件数が実際よりも多くなっている可能性がある。

・ 論文番号は「資料編 1. 論文リスト」に対応している。

・ *：研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの。

3. 学会招待講演・基調講演

1. “Biosynthetic mechanism prediction for a secondary metabolite, ustiloxin B, in *Aspergillus flavus*”, Symposium for Protein bioinformatics at the 4th Asia Pacific Protein Association (APPA) Conference, Jeju, Korea (韓国・済州島)、2014 年 5 月 20 日予定 (<http://www.appa2014.or.kr/program.html>)
2. “EzCatDB: Enzyme reaction database”, Protein bioinformatics and community resources retreat, organized as part of the Wellcome Trust Scientific Conferences programme, Hinxton, UK, (Invited by Alex Bateman (NAR Database issue のエディタ) and Patsy Babbitt) 2014 年 8 月 11 日～12 日予定 (<https://registration.hinxton.wellcome.ac.uk/register.asp?id=440>)

4. 新聞発表等

該当なし

5. 特許出願・成立

特願 2014-044277 「真菌由来のリボソームペプチドを製造する方法」

6. 学会賞等の受賞

該当なし

7. グラントの獲得実績

1. 補酵素結合様式を考慮した次世代活性部位探索アルゴリズムの開発：日本学術振興会/基盤研究(C) H23～H26

8. 書籍等の執筆実績

該当なし

9. 総説の執筆実績

1. 長野希美、「酵素関連データベースの活用－産業応用可能性の高い酵素蛋白質の検索」、月刊バイオインダストリー、株式会社シーエムシー出版、25 巻 5 号、 53～61 頁 (2008)

10. 参加研究者の活動状況（個人情報が含まれるため非公開）