

ライフサイエンスデータベース統合推進事業  
(統合化推進プログラム 2017年採択課題)

## 研究開発終了報告書

「ゲノム・疾患・医薬品のネットワークデータベース」  
金久 實 | 京都大学 化学研究所 特任教授

研究開発期間：2017年4月～2022年3月



©2022 金久 實(京都大学) Licensed under CC BY 4.0

## §1. 研究開発実施の概要

本研究開発では、ゲノムの情報から疾患や医薬品に関する知見を得るための新しいデータベースとして KEGG NETWORK の構築を行い、これまでの統合化推進プログラムで開発してきた KEGG MEDICUS 疾患・医薬品統合リソース (<https://www.kegg.jp/kegg/medicus/>) の新たな柱とすることを目的とした。KEGG NETWORK は疾患を分子間相互作用・反応ネットワークのゆらぎ状態と関連づけたもので、具体的には疾患に関連したヒト遺伝子バリエーション、ウイルスその他の病原体、環境物質、さらには医薬品などが、シグナル伝達や代謝をはじめとした様々な分子間相互作用・反応ネットワークにどのような影響(ゆらぎ)を与えているか文献等から知識を集約し、ネットワーク要素とそれらの集合であるネットワークバリエーションマップという形で表現したデータベースである。様々なバリエーションネットワーク(ゆらいだネットワーク)を通して疾患を理解し、医薬品の標的探索などにも活用できるリソースとすることを目指した。

本研究では、ネットワークデータベース、ネットワーク解析、KEGG MEDICUS 統合リソースの 3 つの研究開発項目を設定して開発を行った。(1) ネットワークデータベースでは、対象疾患として、がん、神経変性疾患、内分泌代謝疾患、先天性代謝異常症、ウイルス感染症、細菌感染症をとりあげ、1,408 個のネットワーク要素と 133 個のネットワークバリエーションマップを作成した。ネットワークバリエーションマップは KEGG パスウェイマップブラウザのサイドパネルからもリンクされ、パスウェイマップにはない疾患関連の詳細情報を提供している。また同一疾患における異なるゆらぎの比較解析や、異なる疾患における共通ネットワークの比較解析などに有用である。例えばがんと神経変性疾患ではアポトーシスのネットワークバリエーションマップが、同じ経路に対して抑制と活性化という逆のゆらぎから構成されることが顕著な特徴として見られた。(2) ネットワーク解析では、テキスト検索ツールとネットワークバリエーションマップに対するマッピングツールを開発し、それぞれ既存の検索システムと KEGG Mapper に組み込んだ。またネットワークバリエーションマップをカスタマイズするツールも開発した。(3) KEGG MEDICUS 統合リソースでは、疾患情報と医薬品情報の充実を継続的に行い、とくに添付文書に記載された適応症と KEGG DISEASE 疾患エントリの整合性を高め、KEGG DRUG 医薬品エントリの効能・疾患フィールドを充実させた。医薬品については常に日米欧の最新情報を提供し、毎月更新される日米の医薬品添付文書では KEGG の ID による標準化・規格化を継続して行い、「商品一覧」のインターフェースに医薬品グループレベルでの比較を導入して一般利用者に対する利便性を向上させた。

本研究開発により、KEGG MEDICUS は KEGG NETWORK ネットワークデータベース、KEGG DISEASE 疾患データベース、KEGG DRUG 医薬品データベース、それに日本と米国の医薬品添付文書を統合した形となった。これまでの疾患と病因遺伝子、医薬品と標的分子の関連に、分子ネットワークレベルの情報が追加され、KEGG MEDICUS の有用性が高まったと考えている。当初計画はすべて達成したが、新型コロナウイルス感染症を契機として、ウイルスを単にゆらぎ物質の 1 つとみなすだけでなく、そのゆらぎがどのように生成されるのかを、ウイルスゲノムとの関連で知識ベース化する必要性を感じた。このような方向性の追加研究も行うこととした。

## §2. 研究開発実施体制

### 1. 研究グループ

#### (1) 研究代表者グループ

- 担当項目

以下のすべての項目を担当した。

1. ネットワークデータベース
2. ネットワーク解析
3. KEGG MEDICUS 統合リソース

### 2. 有識者会議等

#### (1) 会議概要

#### (2) 委員一覧

#### (3) 開催歴

有識者会議等は設置していない。

### §3. 研究開発の目的、実施内容及び成果

#### 1. 研究開発の背景

統合化推進プログラムの第一期および第二期より支援を受けて開発・提供を行ってきた KEGG MEDICUS は、2016 年度末時点では、独自に構築している KEGG DISEASE 疾患データベースと KEGG DRUG 医薬品データベースに、日本および米国の医薬品添付文書を統合したリソースであった。添付文書という実社会で用いられている情報を統合することで、例えばそこに記載された適応症の情報をもとに KEGG DISEASE/DRUG における疾患の定義や治療薬の記述を見直して、実用的な価値を高めることができた。KEGG MEDICUS は医療従事者や一般社会にとっても有用なリソースとなった。一方ではゲノムと疾患や医薬品との関連づけについては、KEGG DISEASE における疾患と病因遺伝子リストとの関係、KEGG DRUG における医薬品とターゲットとの関係といった具合に、遺伝子(産物)レベルの情報しか含まれていなかった。そこで本研究開発では新たな研究基盤情報として、疾患や医薬品に関連した分子間相互作用・反応ネットワークの情報を蓄積した KEGG NETWORK を開発し、分子ネットワークレベルでの情報を通して、すなわち「遺伝子(産物)-分子ネットワーク-疾患・医薬品」の形で、ゲノムとの関連づけを行うこととした。

KEGG 全体では、ゲノムの情報と細胞・個体レベルの情報は KEGG PATHWAY データベースを通して関連づけが行われている。KEGG PATHWAY と KEGG NETWORK の違いは以下の通りである。KEGG PATHWAY はすべての生物種を対象に「通常の分子ネットワーク」を表現しているのに対し、KEGG NETWORK はヒトに特化し、疾患に関連した「ゆらいだネットワーク」を表現している。具体的には KEGG PATHWAY が KO (KEGG Orthology) をノードとした分子ネットワークとして作成されているのに対し、KEGG NETWORK はヒト遺伝子およびヒト遺伝子バリエーションをノードとした分子ネットワークとして記述される。KEGG PATHWAY にある疾患パスウェイマップでは、疾患に関連したヒト遺伝子バリエーションが赤字で示されているが、これは単に絵としての表現で、そのリンクは通常のヒト遺伝子になっている。KEGG NETWORK は KEGG PATHWAY にはない疾患関連の詳細情報を提供するデータベースと位置づけられる。

KEGG NETWORK のゆらいだネットワークはヒト遺伝子バリエーションに起因するものだけでなく、ウイルスやバクテリアなどの病原体、環境物質によるゆらぎも含まれる。さらには医薬品も分子ネットワークのゆらぎ状態を補正する別のタイプのゆらぎ物質とみなしている。このようにヒト疾患に関連した様々な要因を統合的に扱う方法論として分子ネットワークが有効であると考えている。一方、新型コロナウイルス感染症による社会的な背景状況の変化により、ウイルスについて新たな視点が必要となった。すなわちウイルスを単にゆらぎ物質の 1 つとみなすだけでなく、それがどのように生成されるのか、ウイルスゲノムとの関連を探るといった視点である。ヒトに感染するウイルスについて、このような方向性の追加研究も行うこととした。

## 2. 研究開発対象のデータベース・ツール

### (1) データベース

- 主要なもの

正式名称	略称	概要
KEGG MEDICUS		ゲノムの情報と疾患との関連および医薬品の作用・副作用との関連が、生体システムを構成するネットワーク情報として統合され、ゲノムシーケンシングデータ等の大量データの解釈と有効利用を促進するリソースである。また日本と米国のすべての医薬品添付文書も統合されており、一般社会にとっても有用なリソースとなっている。

- 上記以外のもの

正式名称	略称	概要
------	----	----

### (2) ツール等

正式名称	略称	概要
医薬品相互作用チェック		与えられた医薬品リストの中に併用禁忌・併用注意に該当する相互作用があるかを判定するツール。医薬品添付文書に記載された相互作用をすべて抽出し、KEGG DRUG/DGROUP で標準化したデータセットを用いている。

※データベース、ツールの詳細は別紙参照。

## 3. 達成目標及び実施計画

### (1) 当初の実実施計画・達成目標

本研究開発では、ゲノムの情報から疾患や医薬品に関する知見を得るための新しいデータベースとして、ヒトゲノムのバリエーション(多様性)を生体システムを構成するネットワーク要素のバリエーションとして蓄積した KEGG NETWORK を開発する。ネットワーク要素とは、シグナル伝達ネットワーク、代謝ネットワークをはじめとした分子間相互作用・反応ネットワークの中で、基本的な単位となる部分を定義したものである。KEGG MEDICUS には疾患情報(KEGG DISEASE)、医薬品情報(KEGG DRUG)、ネットワーク情報(KEGG NETWORK)が統合され、臨床的シーケンシングデータの解釈や医薬品標的分子の探索など、ゲノム情報有効利用のための新たなリファレンスリソースとして提供する。実施計画では「ネットワークデータベース」と「ネットワーク解析」の新規開発項目、および「KEGG MEDICUS 統合リソース」の継続開発項目を設定し、以下の通り研究開発を進める。

#### 1. ネットワークデータベース

##### 1-1 ネットワーク要素

KEGG NETWORK では分子間のつながりを表現したネットワーク要素を基本的なエン트리(N 番号エン트리)とし、KEGG MODULE の M 番号エントリで用いた論理表現を拡張した形で定義する。ネットワーク要素のノードはヒト遺伝子、そのバリエーション、ウイルスその他の病原体、環境物質、医薬品などである。モジュールと同じように、AND (+) で複合体を、OR (,) でパラログなどを表現できるようにし、モジュールにはないノード間のつながり(エッジ)のタイプも

記述できるようにする。例えば **Activation, Inhibition, Missing interaction** といったタイプである。これまで **KEGG DISEASE** に蓄積してきた疾患と遺伝子の関係、および **KEGG DRUG** に蓄積してきた医薬品と標的分子の関係の中から、**KEGG** パスウェイマップと対応づけられるものを選別し、そこからネットワーク要素を定義していく。

## 1-2 遺伝子バリエント

遺伝子バリエントについてはすでに多くのデータベースで詳細な情報が蓄積されている。本研究では疾患や医薬品に関与するバリエントのみ考慮し、また1つ1つのバリエントを区別するのではなく、ネットワーク中のノードとして適切になるようグループ化を行う。**KEGG** パスウェイマップのノードが機能オーソログ **KO** でグループ化されていること、統合化推進プログラム第二期の研究で医薬品相互作用ネットワークのノードに医薬品グループ **DG** を導入したことに通じる考え方である。**KEGG GENES** では例えばヒトの **ALK** 遺伝子には **hsa:238** という ID がついており、そのバリエントは **238v1, 238v2** といった形の ID で区別し、**KEGG NETWORK** に付随した **VARIANT** データベースに蓄積する。

## 2. ネットワーク解析

### 2-1 ネットワークエントリ検索

**KEGG NETWORK** を用いたネットワーク解析のツールとして、本研究では2つを開発する。1つはネットワークエントリのテキスト検索ツールで、与えられたキーワードやノード ID (遺伝子 ID など) でヒットしたエントリを表示する機能を、**KEGG** で使用している **DBGET** 検索システムと、リレーショナルデータベースを用いた **KEGG MEDICUS** 検索システムに組み込む。

### 2-2 ネットワークマッピング

もう1つは多数のデータの一括処理で、ネットワークマッピングと可視化のツールを開発する。従来からの **KEGG** パスウェイマッピングでは、ゲノムやメタボロームなどの大量データを **KEGG PATHWAY** データベースにマッピングし、再構築できた(もとのデータと対応づけられた)部分をもとに、例えばどのような代謝経路やシグナリング経路があるかを判定し、大量データの意味づけを行う。ネットワークマッピングでも同様に大量データの解釈を支援し、とくにバリエントを含むゲノムのデータから疾患や医薬品に関する知見が得られるようにする。

## 3. KEGG MEDICUS 統合リソース

### 3-1 疾患情報

**KEGG MEDICUS** には上記のネットワーク情報(**KEGG NETWORK**)が追加され、既存の疾患情報(**KEGG DISEASE**)、医薬品情報(**KEGG DRUG/DGROUP**)と統合したリソースとする。疾患情報については、第二期の開発で医薬品添付文書に記載された適応症との対応づけが明確になるような形で **KEGG DISEASE** 疾患エントリの全体的な見直しを行ったが、日本だけでなく米国の医薬品添付文書も用いてこれを継続する。また疾患情報とネットワーク情報の対応づけを行い、ネットワークマッピングなどで疾患に関する知見が得られるようにする。

### 3-2 医薬品情報

医薬品情報については、今後とも日本医薬情報センター(JAPIC)から毎月提供を受けている我が国の医薬品添付文書を **KEGG** の **D/DG/C** 番号で標準化し、**KEGG DRUG** と統合し

て公開する。米国の医薬品添付文書についても同様の標準化を行い、現状通り DailyMed データベースへのリンクとして統合する。また上記の通り日米の医薬品添付文書から適応症を抽出し、KEGG DRUG エントリの DISEASE フィールドに記載する作業を継続する。

### 3-3 統合インターフェース

新規開発するネットワークエントリ検索システムを既存の KEGG MEDICUS 検索システムに組み込み、KEGG DISEASE, KEGG DRUG などと統合的に検索できるインターフェースとする。また添付文書から抽出・標準化したデータセットに基づく医薬品成分(有効成分と添加物)検索ツールと医薬品相互作用チェックのツールの提供も継続する。

KEGG NETWORK に対するネットワークマッピングのツールは KEGG Mapper に組み込んで提供する。KEGG MEDICUS API の開発も継続する。

達成目標としては、研究開始後 1 年以内に KEGG NETWORK データベースの最初のバージョンを公開する。第 3 年次末までに、ネットワーク要素についてノード間の様々な制御関係を定性的に記述できる論理表現法を確立してデータ入力を行い、1,000 個のネットワーク要素をデータベース化する(数値目標は提案書に記載)。またネットワークマッピングツールと可視化ツールを開発し提供する。第 5 年次までには、クリニカルシーケンシングその他の実験データ解釈のために、ネットワークデータベースおよびネットワークマッピングツールを実用的価値のあるものへとする。またゲノム情報を利用した創薬研究、とくに遺伝子バリエーションだけでなくネットワークバリエーションを標的とした医薬品開発の研究という観点で、実用的な価値のあるネットワークデータベースとする。

## (2) 期間中に追加・削除・変更した実施計画・達成目標

主な変更点は 2 つある。第 1 はヒト遺伝子バリエーションの取扱いである。当初計画では dbSNP、ClinVar、COSMIC などの外部のバリエーションデータベースとのリンクづけにより、それらの ID が KEGG NETWORK への入口になると考えていた。しかしながら、外部データベースは著者からの登録に依存したレポジトリであるため、論文等の文献情報から蓄積している KEGG のバリエーションデータとの整合性が問題となった。外部データベースには論文のデータが登録されていなかったり、対応すると思われるハイスループットデータが大量に登録されていてどれを選ぶべきかが分からない、といったことがしばしばあることが明らかとなった。そこで KEGG 独自の遺伝子バリエーション ID が直接利用されるような形で開発を進めていくこととした。

第 2 は当初計画では明確に記載していなかった「ネットワークバリエーションマップ」と呼ぶグラフィカルマップ(nt 番号エントリ)を KEGG NETWORK の主要なデータセットと位置づけたことである。ネットワークバリエーションマップはレファレンスとなる通常のネットワーク要素に対して、関連する様々なゆらぎのネットワーク要素をアライメント表示したものである。現在公開されている KEGG パスウェイマップのブラウザではこの見方に基づき、パスウェイマップ上の通常のネットワーク要素にネットワークバリエーションマップがリンクされ、それを構成する各ネットワーク要素とともに閲覧できるようになっている。

また小規模な追加計画として、ウイルスと宿主細胞の相互作用を理解するために、ウイルスゲノムとウイルスによるヒトの相互作用ネットワークへのゆらぎを関連づける可能性について検討を行うこととした。

## 4. 実施内容

### (1) 実施内容

本研究開発により KEGG MEDICUS はネットワーク情報 KEGG NETWORK(ヒト遺伝子バリエーション情報 VARIANT を含む)、疾患情報 KEGG DISEASE、医薬品情報 KEGG DRUG (DGROUP を含む)、それに外部から導入している医薬品添付文書情報を統合したリソースとなった。研究開発を開始した 2017 年 4 月以降の KEGG MEDICUS のデータ数を表 1 に示した。

表 1. KEGG MEDICUS データ数の推移

	17/4/1	18/4/1	19/4/1	20/4/1	21/4/1	21/10/1
NETWORK(N)		296	690	1,011	1,312	1,408
NETWORK(nt)			88	114	128	133
VARIANT		135	245	415	441	455
DISEASE	1,773	2,036	2,298	2,402	2,498	2,507
DRUG	10,440	10,506	10,955	11,240	11,448	11,735
DGROUP	2,000	2,052	2,206	2,274	2,318	2,382

(N) - ネットワーク要素

(nt) - ネットワークバリエーションマップ

#### ① ネットワークデータベース

ネットワークデータベースについて最初に行ったのは、データを登録・編集するためのデータベース管理システムの開発である。2017 年 5 月末までにデータベースのデザインを決定し、それまでの KEGG データベースの管理システムを拡張する形で NETWORK と VARIANT の入力編集システムを稼働させた。その後、内部用のデータベース管理システムから外部公開のためのフラットファイルデータベースを出力できるようにし、NETWORK と VARIANT を DBGET 検索システムに組み込んだ。2017 年 6 月から田辺麻央が中心となってデータ入力を開始した。まず 17 種類のがんについて論文をもとに、ヒト遺伝子バリエーションによるシグナル伝達と転写制御のバリエーションネットワーク(ゆらいだネットワーク)を定義し、6 種類のがんウイルスについてもウイルスタンパク質によるバリエーションネットワークを定義してデータベース化を行った。ヒトおよびウイルスのタンパク質については古道美穂が KO (KEGG Orthology) の定義などを行っている。

がんを対象とした KEGG NETWORK の最初のバージョンは、達成目標で 1 年以内とした通り、2017 年 12 月 1 日に KEGG MEDICUS (<https://www.kegg.jp/kegg/medicus/>) の一部として公開した。その後はがんウイルス以外のウイルス感染症、一部の細菌感染症、内分泌代謝疾患と先天性代謝異常症、神経変性疾患といった具合に対象疾患を拡大してきた。第 3 年次末までに 1,000 個のネットワーク要素を公開し(表 1)、当初計画の数値目標を達成した。がんについては最終年度に Chemical carcinogenesis を対象にし、化学物質によるゆらいだネットワークも追加した。ネットワーク要素の内訳は表 2 の通りである。



表 2. ネットワーク要素の数

Type			
Reference network		312	レファレンスとなるネットワーク (パスウェイマップに対応)
Variant (perturbed) network	Human gene variant	525	ヒト遺伝子バリエーション
	Pathogen (virus)	370	ウイルスまたは細菌のタンパク質
	Pathogen (bacteria)	104	
Environmental factor		70	化学物質などの環境因子
Drug-target relation		27	ヒト遺伝子バリエーションをターゲットまたはマーカーとする医薬品
Total		1,408	

ネットワークバリエーションマップはレファレンスとなるネットワーク要素に対して、様々なバリエーションネットワーク(ゆらいだネットワーク)をアライメント表示したものである。KEGG NETWORKの最初のバージョンとともに、グラフィカルな表示法として公開し、第2年次以降はあらかじめ定義したバリエーションマップに nt 番号をつけ、データベースのエントリとして管理することとした。図1は神経変性疾患におけるカルシウムシグナリング(nt06410)の一部で、遺伝子バリエーション(赤字)によりどのような分子ネットワークが活性化されるか、AD(アルツハイマー病)、PD(パーキンソン病)、ALS(筋萎縮性側索硬化症)、HD(ハンチントン病)、SCA(脊髄小脳失調症)、PrD(プリオン病)との対応とともに示されている。ネットワークバリエーションマップの表記法として、色付けやエッジの意味は図2に示した通りである。

N00984		Glutamate	→	GRM5	→	GNAQ	→	PLCB	→	IP3	→	ITPR	→	Ca2+	-	MCU		
N01002	AD	APP*	→	Abeta	→	GRM5	→	GNAQ	→	PLCB	→	IP3	→	ITPR	→	Ca2+	-	MCU
N01151	ALS											SIGMAR1*	→	ITPR	→	Ca2+	-	MCU
N01008	AD											PSEN1*	→	ITPR	→	Ca2+	-	MCU
N01007	AD											(PSEN1*,PSEN..	→	RYR3	→	Ca2+	-	MCU
N01199	PrD											PRNP*	→	(RYR,ITPR)	→	Ca2+	-	MCU
N00985	HD			Glutamate	→	GRM5	→	GNAQ	→	PLCB	→	IP3	→	(ITPR+HTT*)	→	Ca2+	-	MCU
N00957	SCA			Glutamate	→	GRM1	→	GNAQ	→	PLCB	→	IP3	→	(ITPR1+(ATXN..	→	Ca2+	-	MCU

図 1. ネットワークバリエーションマップの例(nt06410の一部)

Network element	Coloring	Edge	Signaling network	Metabolic network
Reference network	Green	→	Activation	Enzyme-product relation
Variant network containing	Human gene variant	⊣	Inhibition	Substrate-enzyme relation
	Pathogen gene/protein	—	Complex formation	
	Environmental factor	Purple	≠	Missing interaction
Drug-target relation	Navy	⇒	Expression	Enzyme-enzyme relation
		⇐	Repression	
		⇔		

図 2. ネットワークバリエーションマップの表記法

ネットワークバリエーションマップはどのような観点で関連したネットワークの集合を作るかで2つのタイプがある。1つはシグナリングその他のパスウェイごとに異なる疾患(ただしがん

や神経変性疾患といった同一疾患カテゴリ内の)に関連したネットワークを集めた Pathway view、もう 1 つは疾患ごとに異なるパスウェイを一覧表示した Disease view である。図 1 は Pathway view の例であった。ネットワークバリエーションマップは後述するように、利用者が N 番号とアライメントのためのインデント数を定義することで自由に作成できるが、デフォルトで提供されているものは表 3 の通りである。

表 3. ネットワークバリエーションマップの数

Type	Disease category		
Pathway view	96	Cancer	28
		Neurodegenerative disease	16
		Endocrine and metabolic diseases	9
		Congenital metabolic disorder	17
		Viral and bacterial infection	26
Disease view	37	Cancer	17
		Neurodegenerative disease	6
		Endocrine and metabolic diseases	2
		Viral infection	12
Total	133		

ネットワーク要素のノードとなるヒト遺伝子バリエントは、これまでの KEGG にはないオブジェクトであり、新たに VARIANT データベースに蓄積している。医薬品のターゲットあるいはマーカーとなるバリエントが N 番号エンタリで存在する場合は、表 2 にある Drug-target relation として別の N 番号エンタリを定義して、ネットワークバリエーションマップに表示されるようになっている。図 3 はその例で、BRAF 変異を直接ターゲットにするか、BRAF 変異がある場合に (BRAF 変異をマーカーとして) その下流の MEK をターゲットとする医薬品が表示されている。表 4 は添付文書をもとにバリエントが関与する医薬品をまとめた Brite table ファイル Pharmacogenomic biomarkers (br08341) の内容で、これをもとにネットワークを定義する作業も行っている。

表 4. 遺伝子バリエントを標的またはマーカーにした医薬品の数

がんの分子標的治療薬	91
先天性代謝異常症の酵素補充薬等	29
その他の遺伝子治療薬 (CF、DMD)	9
病原体ゲノムバリエントに対する抗微生物薬	1
医薬品応答に影響を与える多型と変異	135
計	265

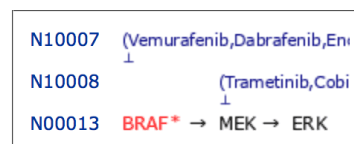


図 3. 医薬品のネットワーク

ネットワークバリエーションマップでは図 1 に示したように、N 番号の後に疾患名(がんの場合はカラーコーディングで区別)が表記され、同一疾患カテゴリ内(ここでは神経変性疾患)で異なる疾患を比較することができるようになっている。一方、疾患カテゴリごとの比較も可能で、例えばアポトーシスのネットワークバリエーションマップにはがんに対応した nt06231 と神経変性疾患に対応した nt06414 があり、同じ経路に対して、がんではアポトーシスを抑制し、神経変性疾患ではアポトーシスを活性化するバリエーションから構成されていることが分かる。このように KEGG NETWORK は疾患と分子ネットワークの関連づけに有用なリソースとなったと考えている。

## ② ネットワーク解析

ネットワーク解析のツールとして、テキスト検索、ネットワーク可視化、ネットワークマッピングのツールを開発した。まずテキスト検索では既存の 2 つの検索システムに組み込む形で実現した。1 つは NETWORK と VARIANT のフラットファイルデータベースを DBGET 検索システムに組み込んで検索できるようにした。NETWORK データベースではネットワーク要素の N 番号エン트리とともに、ネットワークバリエーションマップの nt 番号エントリのテキスト情報も検索できるようになっている。もう 1 つは KEGG MEDICUS 検索システムへの組み込みで、これはフラットファイルではなくリレーショナルデータベースを直接検索する。KEGG NETWORK のページでは NETWORK または VARIANT に対してキーワード入力をする形になっているが、実際は統合的な検索がなされ、キーワードがヒットした DISEASE、DRUG、添付文書などのエン트리も同時に分かるようになっている。

ネットワーク可視化とは図 1 に示したネットワークバリエーションマップのインターフェースで、これを自動的に生成するツールを開発した。あらかじめ定義されたネットワークバリエーションマップは nt 番号を指定することで、あるいはアライメントのためのインデント数を含めた任意の N 番号セットに対しても表示できる。そのために利用者が N 番号セットを編集できるツールも開発した。これはネットワークバリエーションマップのメニューにある Edit Network で起動され、利用者が並べ替えやアライメントなどの編集をして、ネットワークバリエーションマップをカスタマイズできるようになっている。N 番号セットデータの入出力機能もあり、ファイルで編集・保存することも可能である。

ネットワークマッピングとはネットワークバリエーションマップをターゲットデータベースとし、ヒト遺伝子、そのバリエーション、化合物などのデータセットからどのような生体システムの分子ネットワークが関連しているか、またどの遺伝子バリエーションがどのようにネットワークにゆらぎを与え、疾患に関連しているか、といった解析を行う方法論で、そのツールを開発し KEGG Mapper の一部として公開した。5,7) 具体的には KEGG Mapper の Search tool で Search mode を hsa(ヒト)にし、ヒト遺伝子 ID、ヒト遺伝子バリエーション ID、および化合物等の C/G/D 番号のデータセットを入力すると、PATHWAY、BRITE、MODULE、NETWORK、DISEASE の 5 つのターゲットデータベースが一括検索され、それぞれの結果は別のタブに分けて表示されるようになっている。

## ③ KEGG MEDICUS 統合リソース

統合化推進プログラムの第一期より継続して開発してきた KEGG MEDICUS 統合リソースでは、医薬品添付文書という実社会で用いられているデータを統合することで、KEGG DISEASE 疾患データベースおよび KEGG DRUG 医薬品データベースの質が大きく向上した。とく

に本研究開発では、添付文書に記載された適応症にあう形で疾患の定義を見直して疾患エントリの追加・更新を行い、KEGG DRUG の疾患フィールドに適応症の H 番号を入力すると、対応する KEGG DISEASE の医薬品フィールドは逆引きの D 番号が自動生成されるようにした。以前のように文献ベースで集めた治療薬ではなく、実際に使われている治療薬に限定しているのである。

KEGG DISEASE のその他の主な改善点として、疾患名に Subgroup(下位グループ)と Supergroup(上位グループ)の階層を導入し、疾患エントリの分類レベルと相互の関係が明確になるようにした。これに関連して、KEGG DISEASE にある病因遺伝子のリストを KEGG Mapper の Reconstruct ツールに入力し、MODULE マッピングの結果をみると、先天性代謝異常症の病因遺伝子群が多数の complete module を構成することが分かる。例えば尿素サイクル異常症という上位グループの疾患名は、遺伝子ごとの複数の酵素欠損症から構成されているが、その遺伝子群が代謝系の機能モジュールに対応するのである。疾患と機能モジュールの相関として興味深い。5)

KEGG DRUG 医薬品データベースでは、日米欧の新薬を常時チェックして、最新のデータが含まれている。本研究開発では、とくに医薬品と分子ネットワークとの関連で重要なターゲット分子と薬物代謝の情報を整備した。日本医薬情報センター(JAPIC)から提供を受けている医薬品添付文書には独自の付加価値をつけている。とくに「商品一覧」のインターフェースとして、有効成分(D 番号)が一致するもの、有効成分のケミカルグループ(Chemical タイプの DG 番号)が一致するもの、有効成分のクラス(Class タイプの DG 番号)が一致するもの、さらにそれぞれで薬効分類番号も一致するものなどを提供している。また併用禁忌と併用注意に伴う相互作用をすべての添付文書から網羅的に抽出し、KEGG の ID で標準化・規格化を行うことにより、医薬品相互作用チェックツールなどを提供することが可能となっている。

統合インターフェースとしては、英語版の KEGG MEDICUS 検索システムに NETWORK と VARIANT を追加し、KEGG API でも同様の拡張を行った。

#### ④ 追加計画:ウイルス-宿主細胞相互作用

KEGG NETWORK では、ヒト細胞での分子間相互作用・反応ネットワークをネットワーク要素として定義しており、ウイルスタンパク質によるゆらぎもそのバリエーションの 1 つである。ただヒト細胞内のネットワークの詳細が実験的に調べられているウイルスは限られており、ここではウイルスとヒト細胞(あるいはより広く他の宿主細胞)の最初の相互作用に着目して、どの程度の実験データがあり、それがゲノムのデータとどの程度関連づけられるかを調べることにした。具体的にはエンベロープウイルスの細胞侵入に関与する分子間相互作用データで、50 ほどのウイルスのエントリタンパク質とヒタンパク質の相互作用データを文献から集積し、Brite table ファイル(br03220)として公開した。現在の KEGG NETWORK にはヒト細胞内の Chemokine signaling(nt06124)とともに、最初の相互作用だけの Cytokine-cytokine receptor interaction(nt06150)に関するウイルスのネットワークバリエーションマップがある。下流のネットワークを含まない相互作用データなら、すなわちウイルスのタンパク質がヒトのどのタンパク質と相互作用するかのだけであれば、かなりの実験データが存在するようであるが、これをどのように表現するか(br03220 の形か、nt06150 の形か、さらに別の形か)は今後の検討課題である。

KEGG NETWORK ではウイルスや細菌のタンパク質は KO(KEGG Orthology)で表現され、類似のタンパク質への一般化ができるようになっている。KO はこれまで遺伝子単位に

定義され、タンパク質レベルのプロセッシングは区別しない立場をとってきたが、ウイルスでは切断されたペプチド鎖に機能を付与しなければならない場合が多いので、KEGG GENES に新たに vp (viral peptide) のカテゴリを導入した。

## (2) 実施内容のうちの特定項目の詳細

### ① 研究コミュニティを含むデータ提供者との連携・協業

KEGG NETWORK/DISEASE/DRUG を含め、KEGG のデータはすべて文献等から内部メンバーが独自に入力しており、データ提供者は存在しない。ゲノム配列データについても、独自に選別したものを NCBI 等から取得している。ただし、研究コミュニティとはウェブフォームを通じた Feedback system でつながっており、論文に報告されたデータであれば、自分のデータをパスウェイマップに追加してほしい、自分たちが決めたゲノムを KEGG に登録してほしいといった要望には積極的に応えている。

### ② データベース利用者への周知、利用者との連携・協業

KEGG NETWORK を含め、新規に公開した内容については 1~3 ヶ月ごとに Release notes としてウェブで公開している。また Nucleic Acids Research 誌の Database Issue で 2 年ごとに KEGG の最新状況を報告している。3,6) ネットワークマッピングを含めた KEGG Mapper については Protein Science で論文発表を 2 回行った。5,7)

データベース利用者との間も Feedback system でつながっており、様々な質問や有用なコメントをもらい、データベースの質の向上につながっている。

### ③ 利用者にとって有用なデータ基盤の構築

KEGG MEDICUS では 2 つのタイプのデータ統合を行っている。1 つは医薬品添付文書という実社会で使われているデータを研究基盤としての KEGG DRUG/DISEASE データベースに取り込む形での統合で、KEGG MEDICUS の実用的価値が高まった。もう 1 つは DB 間連携のための統合で、ヒト遺伝子バリエーション、疾患、医薬品などで、対応する様々な外部データベースへのリンクづけを行い、ゲノムネットの LinkDB システムを通して RDF データでの提供を行っている。

データ利活用を促進するための技術開発としては KEGG Mapper の改良を行い、KEGG NETWORK のネットワークバリエーションマップを組み込んだ。

### ④ 持続的なデータベース運用体制の構築に向けた取り組み

KEGG は人手で構築しているデータベースであるが、データ入力ツールに付随したチェック機能とともに、入力後のデータについては他のデータも含めた様々なチェックプログラムが 1 時間ごとに問題点を指摘するようになっている。

KEGG MEDICUS を含め KEGG のデータは毎日更新されているが、利用ツールも含めた機能的な改善等は、記録を残す意味もあり 1~3 ヶ月ごとに公開して Release notes で案内している。

### ⑤ 人材の育成

普通の意味での人材の育成はとくに実施していないが、上記のチェックプログラムを導入してからは、自分のミスを見出しやすくなり、また自然とアノテーションの質が向上した。

KEGG の研究開発者は長期間在籍する方が多く、キャリアパス支援は行っていない。

⑥ 国際連携・国際貢献

KEGG MEDICUS は日本の医薬品添付文書を提供しているため国内からの利用者が圧倒的に多い(海外は1~2割しかない)が、KEGG 全体では海外からの利用者が7割程度である。国際的にユニークな高品質のデータベースとして定着しており、アクセス数や論文引用回数からも国際貢献度は高いと考えている。

⑦ その他

とくにありません。

## §4. 主要なデータベースの利活用状況

### 1. アクセス数

#### (1) 実績

表 1 研究開発対象の主要なデータベースの利用状況(月間平均)

名称	種別	2017年度	2018年度	2019年度	2020年度	2021年度 (8月時点)
KEGG MEDICUS	訪問者数	695,248	1,341,147	2,425,518	1,956,422	1,583,110
	訪問数	1,205,103	2,321,565	4,691,379	3,668,175	2,849,595
	ページ数 (海外分)	2,117,312 (237,315)	3,733,736 (304,604)	7,657,800 (443,279)	6,519,650 (618,259)	5,062,571 (690,037)

#### (2) 分析

- ここに示したのは KEGG MEDICUS へのアクセス(KEGG 全体ではない)である。研究開始以前の 2016 年度は訪問者数が 252,491、訪問数が 419,254、ページ数が 892,986 で、2019 年度にはいずれも 10 倍になっている。これは Google 等の検索エンジンから医薬品添付文書のページに入るトラフィックが急増したためである。とくに 2018 年 9 月のクラリベイト・アナリティクス引用栄誉賞により KEGG の認知度が高まったことで KEGG MEDICUS 医薬品添付文書のページのランキングも高くなったようである。
- その後は Google での医薬品添付文書ページのランキングが多少下がっており、他の競合するページが目立つようになっている。これはコロナウイルス感染症により医薬品情報に対する社会的ニーズの高まりを反映しているのかもしれない。
- ただし、海外からのアクセスに限定すると、アクセスページ数は毎年増加している。これは日本語の医薬品添付文書以外の部分、すなわち独自に構築している KEGG MEDICUS の部分に対応すると考えられる。

### 2. データベースの利用状況を示すアクセス数以外の指標

以下は KEGG 全体の利用状況である。

発表論文に対する引用件数

発表論文	DOI	出版年月日	引用件数
Nucleic Acids Research 2019 <sup>3)</sup>	10.1093/nar/gky962	2018.10.13	766
Protein Science 2019 <sup>4)</sup>	10.1002/pro.3715	2019.8.22	385
Protein Science 2020 <sup>5)</sup>	10.1002/pro.3711	2019.8.18	209
Nucleic Acids Research 2021 <sup>6)</sup>	10.1093/nar/gkaa970	2020.10.30	209
Protein Science 2021 <sup>7)</sup>	10.1002/pro.4172	2021.8.23	

引用件数は Web of Science より 2021 年 9 月 29 日に取得

#### 学術雑誌掲載のための利用許諾申請件数(月間平均)

年度	2017年度	2018年度	2019年度	2020年度	2021年度
月間平均件数	29	31	39	88	117

KEGGの利用者が解析結果をパスウェイマップ等の図で表現して論文を執筆し、それを学術雑誌に掲載するには、KEGGから利用許諾を得ていることが出版の条件になっている。利用許諾申請の受付は Feedback と類似のウェブフォームを用いており、申請件数は毎週数十件ある。ここでは1ヶ月あたりの申請件数の平均値を示した。この2年間でとくに中国から、またとくに Scientific Reportsをはじめとした open access journal が急増している。

### 3. データベースの利活用により得られた研究成果（生命科学研究への波及効果）

Nucleic Acids Research Database Issue に出した13報のKEGG論文を引用している論文、すなわちKEGGデータベースを利用したと思われる論文をWeb of Scienceで調査したところ、合計で33,510件で、2017年は2,843件、2018年は3,206件、2019年は4,114件で、2020年は4,720件、2021年は9月までで3,872件であった。2017年以降でNature誌に出た論文は33件あり、その中で医療に関連していると思われるものは以下の10件を含め20件近くあった。

1. Klunemann M, et al., “Bioaccumulation of therapeutic drugs by human gut bacteria”, Nature 597, 533-538, 2021 (DOI:10.1038/s41586-021-03891-8).
2. Han S, et al., “A metabolomics pipeline for the mechanistic interrogation of the gut microbiome”, Nature 595, 415-420, 2021 (DOI:10.1038/s41586-021-03707-9).
3. Lawson KA, et al., “Functional genomic landscape of cancer-intrinsic evasion of killing by T cells”, Nature 586, 120-126, 2020 (DOI:10.1038/s41586-020-2746-2).
4. Cummings BB, et al., “Transcript expression-aware annotation improves rare variant interpretation”, Nature 581, 452-458, 2020(DOI:10.1038/s41586-020-2329-2).
5. Shuai SM, et al., “The U1 spliceosomal RNA is recurrently mutated in multiple cancers”, Nature 574, 712-716, 2019 (DOI:10.1038/s41586-019-1651-z).
6. Chu C, et al., “The microbiota regulate neuronal function and fear extinction learning” Nature 574, 543-548, 2019 (DOI:10.1038/s41586-019-1644-y)
7. Blacher E, et al., “Potential roles of gut microbiome and metabolites in modulating ALS in mice” Nature 572, 474-480, 2019 (DOI:10.1038/s41586-019-1443-5)
8. Johnson EO, et al., “Large-scale chemical-genetics yields new M. tuberculosis inhibitor classes” Nature 571, 71-78, 2019 (DOI:10.1038/s41586-019-1315-z)
9. Lloyd-Price J, et al., “Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases” Nature 569, 655-662, 2019 (DOI:10.1038/s41586-019-1237-9)
10. Jiang Y, et al., “Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma” Nature 567, 257-261, 2019 (DOI:10.1038/s41586-019-0987-8)

### 4. データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション（産業や科学技術への波及効果）

世界のトップ製薬会社やその他の多くの会社はKEGGを活用しているが、何のために使っているのか、どのような結果が得られたのか、といった点については一切把握していない。ただし、活用されているということで産業への波及効果があるものと推定される。



## §5. 今後の展開

KEGG NETWORK はヒトゲノムと疾患との間に分子ネットワークを導入したものであるが、病原体ゲノムと疾患との関連についてもネットワーク的なアプローチが可能であると考えている。KEGG NETWORK ではウイルスおよびバクテリアのタンパク質はすでに KO 化されており、KEGG パスウェイマップのように KO を通して未知のウイルスあるいは未知のバクテリアでネットワーク再構築できる可能性がある。統合化推進プログラム第二期ではバクテリアゲノムと薬剤耐性の関連を調べ、細分化された KO が予測のために有効であることを確認した。一方、ウイルスの KO は一般にごく近縁種でしか保存されておらず、ユニーク性が高い。また現在 KEGG に含まれる生物種では 3560 万のタンパク質の 52%に KO がつけられているのに対し、57 万のウイルスタンパク質のうち KO がつけられたものはわずか 2%しかない。ただ逆の言い方をするとウイルスでは実験的に確かめられたデータや知識を適切に集積すれば、発展の余地が大きいのもかもしれない。その際、相互作用するウイルスタンパク質だけでなく、それがゲノムからどのようなプロセッシングで生成されているか、さらにはウイルスの系統関係およびホストの範囲とその系統関係といったものを考慮する必要があるだろう。このように統合的に解析できる方法論があれば、宿主(**cellular organism**)とウイルスの共進化を理解し、宿主内でのウイルスの変化を予測するといったアプローチができるかもしれない。ウイルスに着目すると KEGG NETWORK の発展の可能性は大きいと考えている。

## §6. 自己評価

KEGG データベースには生命の基本原理を探求するといった基礎的な側面と、ヒト疾患の理解や医薬品開発支援などで社会貢献を行うといった応用的側面がある。後者の側面を担うのが本研究の KEGG MEDICUS である。KEGG NETWORK という全く新規のデータベースを 1 年足らずで公開し、3 年後、5 年後の達成目標に十分到達できたのは KEGG 全体の様々な資産を活用できたからである。統合化推進プログラムの本来の目的は他のグループで開発されているデータベースとの連携・統合化を行うことであるが、本研究開発で行ったことは KEGG 本体との連携・統合化であった。しかしながら本研究開発により、新たな情報基盤を構築するという本プログラムの趣旨には貢献できたと考えている。

KEGG MEDICUS には医薬品添付文書という外部から導入しているデータが含まれ、これがあるために一般社会から膨大なアクセスがあり、KEGG DISEASE/DRUG の品質の向上にも役に立っている。ただ添付文書を提供していることが評価に値するものではなく、今後ともこれを KEGG MEDICUS の一部として維持すべきかどうかは検討の余地がある。現時点では維持する方向ではあるが、現在は KEGG 内部のリレーショナルデータベースの一部となっている添付文書データを、独立したリレーショナルデータベースで管理するようにし、将来的な状況の変化に対応できるようにした。

## §7. 外部発表等

### 1. 原著論文発表

#### (1) 論文数概要

種別	国内外	件数
発行済論文	国内(和文)	1 件
	国際(欧文)	8 件
未発行論文	国内(和文)	0 件
	国際(欧文)	0 件

#### (2) 論文詳細情報

1. Minoru Kanehisa, "Enzyme annotation and metabolic reconstruction using KEGG", In "Protein Function Prediction: Methods and Protocols" (Kihara, D., ed.), Humana Press, Methods in Molecular Biology, vol. 1611, pp. 135-145, 2017 (DOI: 10.1007/978-1-4939-7015-5\_11).
2. Minoru Kanehisa, "Inferring antimicrobial resistance from pathogen genomes in KEGG", In "Data Mining for Systems Biology, Second Edition" (Mamitsuka, H., ed.), Humana Press, Methods in Molecular Biology, vol. 1807, pp. 225-239, 2018 (DOI: 10.1007/978-1-4939-8561-6\_17).
3. Minoru Kanehisa, Yoko Sato, Miho Furumichi, Kanae Morishima and Mao Tanabe, "New approach for understanding genome variations in KEGG", Nucleic Acids Research, vol. 47, No.D1, pp. D590-D595, 2019 (DOI: 10.1093/nar/gky962).  
概要:Nucleic Acids Research の Database Issue で 2 年ごとに出している KEGG 論文で、とくに新しい方法論でゲノムのバリエーションをデータベース化している KEGG NETWORK について報告した。
4. Kanehisa, M; Toward understanding the origin and evolution of cellular organisms. Protein Science, vol. 28, No. 11, pp. 1947-1951, 2019 (DOI: 10.1002/pro.3715).
5. Minoru Kanehisa and Yoko Sato, "KEGG Mapper for inferring cellular functions from protein sequences", Protein Science, vol. 29, No. 1, pp. 28-35, 2020 (DOI: 10.1002/pro.3711).
6. Minoru Kanehisa, Miho Furumichi, Yoko Sato, Mari Ishiguro-Watanabe and Mao Tanabe, "KEGG: integrating viruses and cellular organisms", Nucleic Acids Research, vol. 49, No. D1, pp. D545-D551, 2021 (DOI: 10.1093/nar/gkaa970).  
概要:Nucleic Acids Research の Database Issue で 2 年ごとに出している KEGG 論文の最新版。これまで細胞性生物 (cellular organism) を中心にしてきた KEGG で、ウイルスをどのように扱うか、KEGG NETWORK とともに新しい方向性を示した。
7. Minoru Kanehisa, Yoko Sato and Masayuki Kawashima, "KEGG mapping tools for uncovering hidden features in biological data", Protein Science, vol. 31, No. 1, 2022 (DOI: 10.1002/pro.4172).

#### 2. その他の著作物(総説、書籍など)

1. Minoru Kanehisa, "KEGG GLYCAN", In "A Practical Guide to Using Glycomics Databases" (Aoki-Kinoshita, K.F., ed.), pp. 177-193, Springer, 2017.

2. 金久 實, “KEGG データベース”, バイオサイエンスとインダストリー (B&I), vol. 76, no. 1, pp. 58-62, 2018.

### 3. 国際学会発表及び主要な国内学会発表

#### (1) 概要

種別	国内外	件数
招待講演	国内	4 件
	国際	6 件
口頭発表	国内	4 件
	国際	2 件
ポスター発表	国内	4 件
	国際	0 件

#### (2) 招待講演

##### 〈国内〉

1. 金久 實, KEGG: 創薬・医療への展開、医薬創成情報科学専攻 10 周年記念シンポジウム、京都大学薬学研究科、2017 年7月 3 日
2. 金久 實, ゲノム・疾患・医薬品の統合リソース KEGG MEDICUS の利用法、MedXML コンソーシアム Seagaia Meeting、宮崎、2018 年 5 月 19 日
3. 金久 實, KEGG for molecular network-based understanding of human diseases、日本バイオインフォマティクス学会年会、東京、2019 年 9 月 9 日
4. 金久 實, ゲノムの情報から生命システム再構築へ ~KEGG がつなぐゲノムと社会~、第 239 回生命科学フォーラム、東京、2019 年 12 月 10 日

##### 〈国際〉

1. Minoru Kanehisa, From gene variants to network variants: a new database for understanding diseases and drugs, BioNetVisA Workshop in Basel Computational Biology Conference, Basel, Switzerland, September 12, 2017
2. Minoru Kanehisa, “Genome annotation for inferring phenotypic features”, WDCM GCM2.0 SIGs Meeting, Beijing, China, June 26, 2018
3. Minoru Kanehisa, “From KEGG Orthology to modules and human diseases”, NIB B Conference: Quest for Orthologs, Okazaki, Japan, August 1, 2019
4. Minoru Kanehisa, “Toward understanding the origin and evolution of cellular organisms”, 8th Tokyo Tech International Symposium, Yokohama, Japan, January 8, 2020
5. Minoru Kanehisa, “KEGG as a viral pathway resource”, 10th Aquatic Virus Workshop, Zoom to Kyoto, June 29, 2021
6. Minoru Kanehisa, “Toward understanding the origin and evolution of cellular organisms and viruses”, Protein Society 35th Anniversary Symposium, Zoom to Virtual meeting, July 7, 2021

### (3) 口頭講演

〈国内〉

1. 金久 實、KEGG NETWORK Version 1: がんのネットワークバリエント、トーゴの日シンポジウム 2017、東京、2017 年 10 月 5 日
2. 金久 實、ネットワークによるゲノム・疾患・医薬品の統合、トーゴの日シンポジウム 2018、東京、2018 年 10 月 5 日
3. 金久 實、KEGG NETWORK: 疾患に関連したネットワークバリエントのデータベース、第 41 回日本分子生物学会年会シンポジウム、横浜、2018 年 11 月 28 日
4. 金久 實、新型コロナウイルスのゲノム解析の課題と対策～ゲノムからシステムを解読する情報技術について～、立憲民主党政務調査会、Zoom to 東京、2021 年 5 月 27 日

〈国際〉

1. Minoru Kanehisa, "Enzyme annotation in KEGG", IUPAC-IUBMB Joint Commission on Biochemical Nomenclature, Cambridge, UK, May 31, 2018
2. Minoru Kanehisa, "Toward understanding the origin and evolution of cellular organisms", European Bioinformatics Institute, Hinxton, UK, September 26, 2019

### (4) ポスター発表

〈国内〉

1. 田辺 麻央、金久 實、がんのシグナル伝達ネットワークバリエントのデータベース、トーゴの日シンポジウム 2017、東京、2017 年 10 月 5 日
2. 田辺 麻央、金久 實、ウイルスによるネットワークバリエントのデータベース、トーゴの日シンポジウム 2018、東京、2018 年 10 月 5 日
3. 田辺 麻央、松浦 有里子、金久 實、内分泌代謝疾患のネットワークバリエーションマップ、トーゴの日シンポジウム 2018、東京、2019 年 10 月 5 日
4. 田辺 麻央、金久 實、KEGG NETWORK による疾患関連遺伝子バリエント、ウイルス、環境因子の統合、トーゴの日シンポジウム 2021、Zoom to 東京、2021 年 10 月 5 日

〈国際〉

該当なし

## 4. 知財出願

### (1) 出願件数

種別	件数
特許出願	国内 0 件
	国外 0 件
その他の知的財産出願	0 件

### (2) 一覧

- ① 国内出願
- ② 海外出願
- ③ その他の知的財産権

## 5. 受賞・報道等

### (1) 受賞

1. \* Clarivate Analytics Citation Laureate, Minoru Kanehisa, September 20, 2018
2. \* Protein Society Carl Brändén Award, Minoru Kanehisa, March 18, 2019

### (2) メディア報道

1. 上記 1 に関連した報道多数、その後も毎年秋に報道あり
2. 読売新聞、生命の起源 コンピューターで迫りたい、2018 年 12 月 7 日
3. 日経産業新聞、生命科学に欠かせぬ DB、2020 年 1 月 20 日
4. 南日本新聞こども新聞、鹿児島から世界へ！、2020 年 12 月 10 日

### (3) その他

1. Thomson Reuters Highly Cited Researchers 2018、金久 實
2. Thomson Reuters Highly Cited Researchers 2021、金久 實

## §8. 研究開発期間中の活動

### 1. 進捗ミーティング

KEGG チーム内の進捗ミーティングは毎週行っている。

また外部のソフトウェア開発チームを含めた進捗ミーティングを毎月行っている。

### 2. 主催したワークショップ、シンポジウム、アウトリーチ活動等

年月日	名称	場所	参加人数	目的・概要
2017年12月13～15日	KEGG MEDICUS の展示 デモ	神戸ポートアイランド	150人	ConBio2017(分子生物学会ほか年会)での展示
2018年1月12日	公開シンポジウム「第9回新たな創薬パラダイムの創出」	東京大学医科学研究所1号館講堂	92人	日本発の創薬技術・新薬に関するシンポジウム、NPOバイオインフォマティクスジャパンと連携
2018年11月28～30日	KEGG MEDICUS の展示 デモ	パシフィコ横浜	100人	分子生物学会年会での展示とデモ
2019年12月3～6日	KEGG MEDICUS の展示 デモ	福岡国際会議場	100人	分子生物学会年会での展示とデモ

以上

別紙 研究開発対象のデータベース等

No.	正式名称	別称	概要	URL	公開日	状態	分類	生命科学系データベースアーカイブ	NBDCヒトデータベース	NBDC RDFポータル	関連文献 (論文リストに記載があれば、その番号でも可)
1	KEGG MEDICUS		ゲノムの情報と疾患との関連および医薬品の作用・副作用との関連が、生体システムを構成するネットワーク情報として統合され、ゲノムシークエンスデータ等の大量データの解釈と有効利用を促進するリソースである。また日本と米国のすべての医薬品添付文書も統合されており、一般社会にとっても有用なリソースとなっている。	<a href="https://www.kegg.jp/kegg/medicus/">https://www.kegg.jp/kegg/medicus/</a>	2010年10月1日	継続・発展	データベース等	公開済	対象外	対象外	文献4,6
2	医薬品相互作用チェック		与えられた医薬品リストの中に併用禁忌・併用注意に該当する相互作用があるかを判定するツール。医薬品添付文書に記載された相互作用をすべて抽出し、KEGG DRUG/DGROUPで標準化したデータセットを用いている。	<a href="https://www.kegg.jp/medicus-bin/ddi_manager">https://www.kegg.jp/medicus-bin/ddi_manager</a>	2016年4月1日	継続・発展	ツール等	対象外	対象外	対象外	