

バイオサイエンスデータベースセンター統合化推進プログラム  
「ゲノムとフェノタイプ・疾患・医薬品の統合データベース」

# ゲノムと疾患・医薬品をつなぐ 相互作用ユニットの知識ベース

京都大学化学研究所

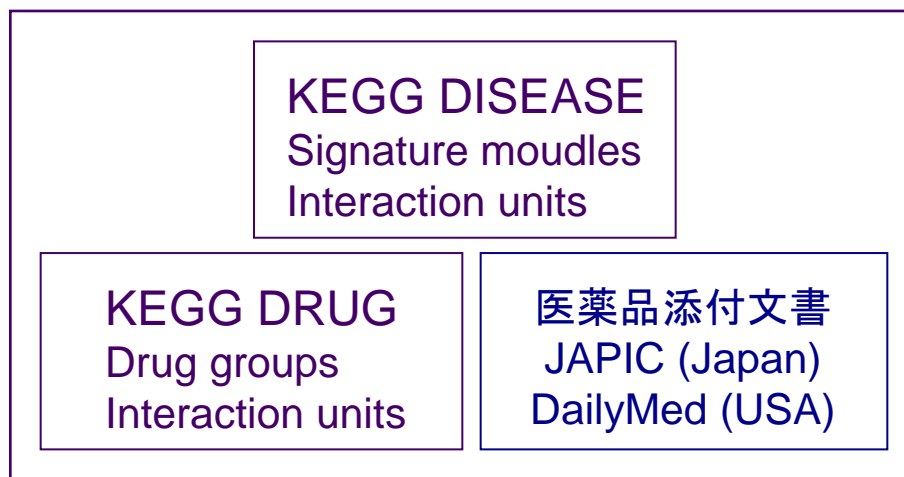
金久 實

2014年10月5日 トーゴーの日シンポジウム

ゲノム情報を有効利用するための  
トランスレーショナルバイオインフォマティクスリソース

# KEGG MEDICUS

ヒトゲノム  
病原体ゲノム  
体内メタゲノム  
環境メタゲノム



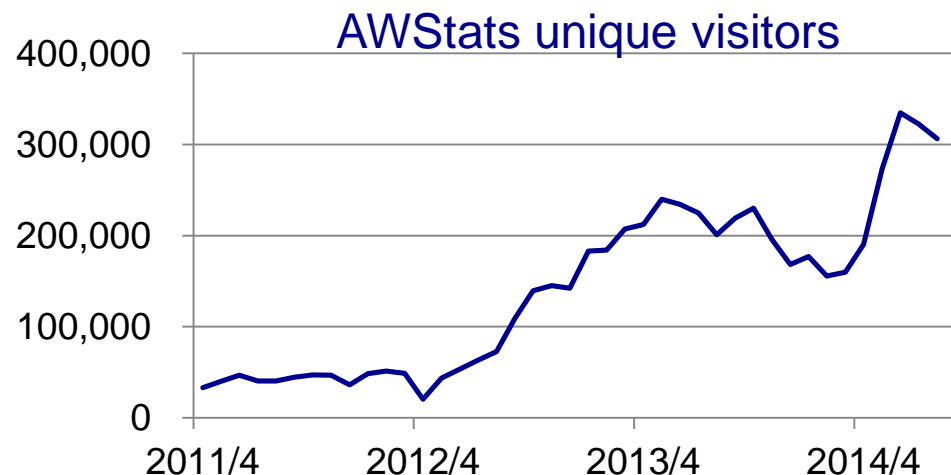
医療・創薬  
研究開発  
実用化



医薬品情報  
サービス

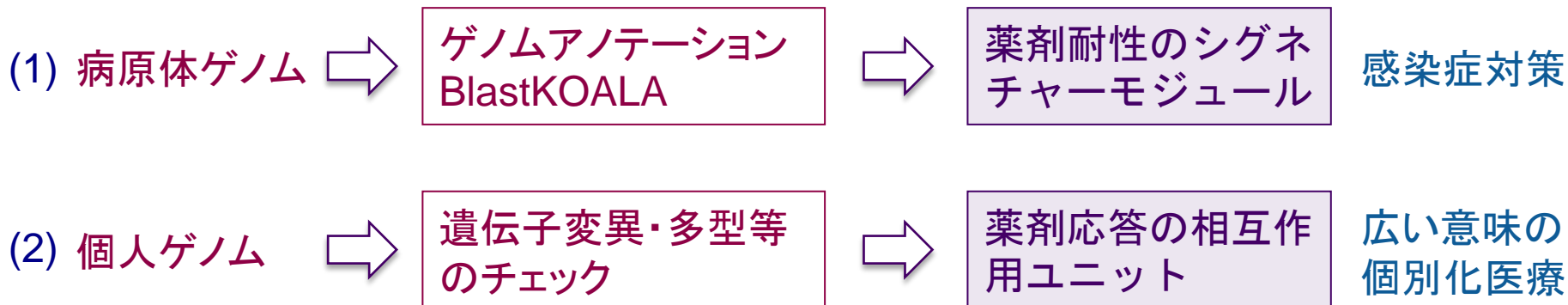
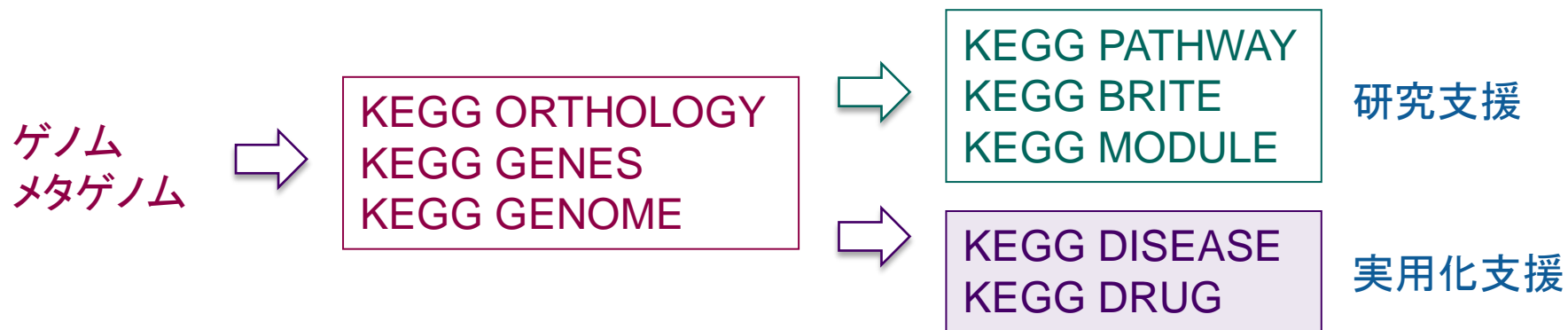
H23-H25: 研究者コミュニティと  
社会との架け橋

H26-H28: 研究者コミュニティの  
トランスレーショナル  
研究を推進



# 研究開発の目標・ねらい

ゲノムとフェノタイプとの関連、とくに疾患との関連を理解するために、個々の遺伝子だけでなく、複数の遺伝子から構成された機能モジュール、さらには遺伝子、タンパク質、環境因子、医薬品等から構成された相互作用ユニットに関する知識をデータベース化する。



# 基本的な考え方

- 限られた実験データから得られる知識をいかに一般化してより多くのデータに適用できるようにするか
- ネットワークでの分子機能 — Network (context) dependent molecular function

---

対象	様々なタイプの分子ネットワーク	生体内化学反応ネットワーク	医薬品相互作用ネットワーク
インスタンス	遺伝子・タンパク質 KEGG GENES	反応 KEGG REACTION	医薬品 KEGG DRUG
クラス	オーソログ KEGG ORTHOLOGY KO	反応クラス KEGG RCLASS RC	医薬品グループ KEGG DGROUP DG
機能単位	KEGG モジュール シグネチャーモジュール	反応モジュール	相互作用ユニット

---

# KEGG ORTHOLOGY (KO) とは

- 分子間相互作用ネットワークのノードとしての同一機能を反映したオースロググループ
- 類似配列のグループでもあるが、類似度の閾値は一定ではない
- KO エントリ（K番号エントリ）には、実験的に明らかにされた遺伝子・タンパク質の機能情報として、論文の PubMed ID と対応する KEGG GENES ID（類似配列グループのコア配列）を蓄積
- 全配列既知のゲノムを網羅的に集積している KEGG GENES に対し人手によるチェックを含む KO アノテーション（K番号づけ）

## BlastKOALA とは

- Nonredundant な KEGG GENES を用いてゲノムやメタゲノムの KO アノテーションを行い、KEGG パスウェイ、BRITE 機能階層、KEGG モジュールを再構築して高次機能解釈を提供するサービス
- 病原体ゲノムについては、シグネチャーモジュールによるフェノタイプ解釈も提供予定

# KEGG モジュールとは

- 分子間相互作用ネットワーク中の機能ユニット
- ゲノム中のユニット（オペロン、プラスミドなど）にしばしば対応
- モジュールエントリ（M番号エントリ）はK番号の論理式で定義
- とくにフェノタイプに関連したユニットをシグネチャーモジュールと呼んでいる

Methicillin resistance  
M00625 = K02547 K02546 K02545

	Organism	1	2	3
		K K02547 (mecR1)[14]	K K02546 (mecI)[14]	K K02545* (mecA)[14]
B.Fir	M sau	SA0039(mecR1)	SA0040(mecI)	SA0038(mecA)
B.Fir	M sav	SAV0042(mecR1)	SAV0043(mecI)	SAV0041(mecA)
B.Fir	M saw	SAHV_0041(mecR1)	SAHV_0042(mecI)	SAHV_0040(mecA)
B.Fir	M sah	SaurJH1_0030	SaurJH1_0031	SaurJH1_0029
B.Fir	M saj	SaurJH9_0030	SaurJH9_0031	SaurJH9_0029
B.Fir	M sar	SAR0040(mecR1)	SAR0041(mecI)	SAR0039(mecA)
B.Fir	M suk	SAA6008_00040(m	SAA6008_00041(m	SAA6008_00039(m
B.Fir	M sut	SAT0131_00035	SAT0131_00036	SAT0131_00034(r
B.Fir	M suf	SARLGA251_00270	SARLGA251_00280	SARLGA251_00260
B.Fir	M sauz	SAZ172_0073(mec	SAZ172_0074(mec	SAZ172_0072(mec
B.Fir	M suy	SA2981_0040(mec	SA2981_0041	SA2981_0039(mec
B.Fir	M sauc	CA347_40(mecR1)	CA347_41(mecI)	CA347_39(mecA)
B.Fir	M ser	SERP2520(mecR1)	SERP2519(mecI)	SERP2521(mecA)
B.Fir	M mcl	MCCL_plsB0022	MCCL_plsB0021	MCCL_plsB0023

D00350 Itraconazole  
K17719 CYP3A4  
DGxxxxx CYP3A4 substrates

DG00149 Warfarin  
K05357 VKORC1  
K17719 CYP2C9

DGxxxxx Statin  
C00187 Cholesterol  
K12473 LDLR mutation

## 相互作用ユニットとは

- モジュールほど明確な機能ユニットではないが、何らかの意味をもつ相互作用の単位
- 遺伝子（およびタンパク質、RNA）だけでなく、化合物、糖鎖、医薬品などを含むユニット
- とくに医薬品間の相互作用ユニットやゲノムと薬剤応答の関連ユニットに着目している

# (1) 薬剤耐性のシグネチャーモジュール

## $\beta$ -ラクタム耐性の例

タイプ	モジュール	遺伝子セット	遺伝子変異
薬剤の標的分子の変化	M00625	mecA – mecR1 – mecl	
	M00626		PBP
薬剤の不活性化（分解）	M00627	penP – blaR1 – blaI	
	M00628	penP – ampR – ampC	
薬剤の透過性低下	M00629		OMP
薬剤の排出	M00649	adeABC – adeR – adeS	

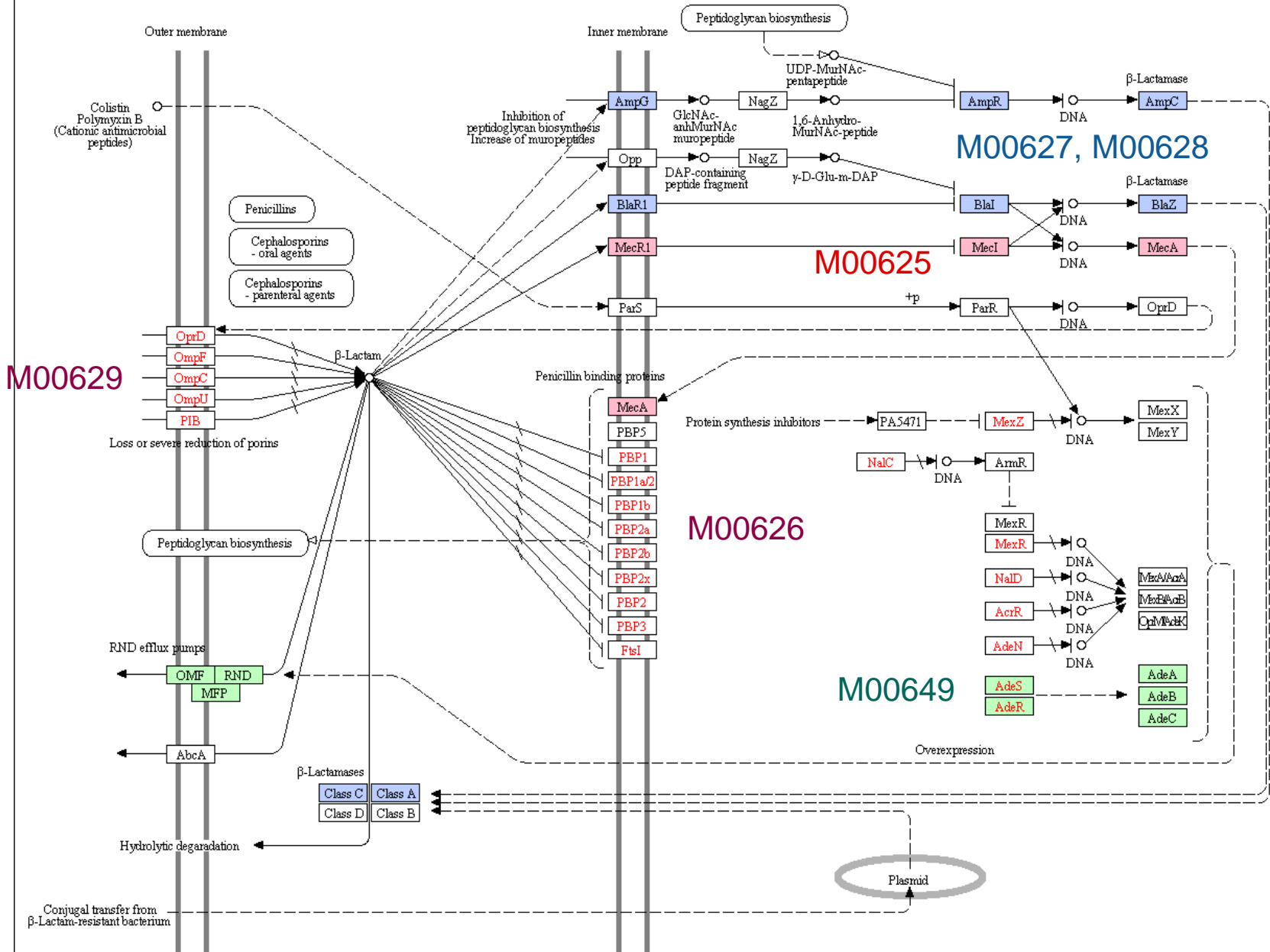
モジュール定義      M00625 = K02547 K02546 K02545

適用可能な生物種グループの検証      Taxonomy checker

適用可能な医薬品グループの検証      Drug group

# β-ラクタム耐性の KEGG パスウェイマップ

beta-LACTAM RESISTANCE







## Taxonomy checker - M00625

Enter single M/K number or logical expression of M/K numbers

complete only  include 1 block missing  include 1 block or 2 block missing

### Definition

M00625 Methicillin resistance

Organisms mapping against:

Grp	Genus	Species	Org
B.Fir	Staphylococcus	Staphylococcus aureus	sau sav saw sah saj sam sar sac sax saa suu suh suj suk sut suq suz sux suw sug suf saun saus sauu sauz suy saub saum sauc saur saui saut sauj sauk sauq sauv sauw saux sauy
		Staphylococcus epidermidis	ser
		Staphylococcus haemolyticus	sha
B.Fir	Macrocooccus	Macrocooccus caseolyticus	mcl

# 病原体ゲノムのモジュール情報は感染症疾患エントリの一部とする

## KEGG DISEASE 疾患エントリの内容

ENTRY	H番号識別子
NAME	疾患名、別名を含む
DESCRIPTION	疾患の簡単な説明文
CATEGORY	疾患カテゴリ
PATHWAY	KEGGパスウェイへのリンク
GENE	既知の病因遺伝子
CARCINOGEN	発癌物質
ENV_FACTOR	その他の環境因子
PATHOGEN	病原体
MODULE	病原性・薬剤耐性のモジュールあるいは遺伝子
MICROBIOME	疾患に関連したマイクロビオームの特徴
MARKER	代表的な診断マーカー
DRUG	代表的な治療薬 (医薬品グループ表記)
COMMENT	コメント
DBLINKS	外部データベースへのリンク
REFERENCE	文献情報

## (2) 薬剤応答の相互作用ユニット

ゲノムバイオマーカー（Pharmacogenomic biomarker）の例

タイプ	医薬品	商品名	マーカー
コンパニオン診断	D01441	グリベック	BCR-ABL 融合遺伝子
	D09731	ザーコリ	EML4-ALK 融合遺伝子
	D03257	ハーセプチン	HER2 過剰発現
	D01977	イレッサ	EGFR 変異
遺伝子多型・変異と薬剤応答	DG00149	ワルファリン	VKORC1,CYP2C9 多型
	DGxxxxx	スタチン	LDLR 欠損
	D00966	タモキシフェン	CYP2D6 多型
	D01155	ゲムシタビン	CDA 多型

添付文書記載のゲノムバイオマーカー  
添付文書にはないが論文等で報告されているもの

# KEGG DGROUP: 医薬品グループデータベース

1. ATC分類最下層に対応した同一基本構造、同一薬効のグループ
2. 添付文書の相互作用欄に記載された同一相互作用のグループ
3. 米国の商品を分類したFDA EPCとVA NDF-RTにあるグループ
4. USP Dictionaryに記載された医薬品名のsuffixに対応したグループ

1 は公開済み



DGROUP: DG00149

Help

エントリ	DG00149	Structure DGroup
名称	ワルファリン	
メンバー	D08682 ワルファリン D00564 ワルファリンナトリウム <US> D01280 ワルファリンカリウム (JP16) <JP>	日本、米国の商品があるものには フラグがついている
コード	ATC code: B01AA03	
コメント	ビタミンK拮抗系抗凝固薬, クマリン系抗凝固薬 標的分子: vitamin-K-epoxide reductase inhibitor [HSA:79001] [KO:K05357]; phylloquinone reductase inhibitor [HSA:1728] [KO:K00355] 代謝酵素: CYP2C9 [HSA:1559], CYP1A2 [HSA:1544]	
LinkDB	<a href="#">All DBs</a>	

2のグループは添付文書から医薬品相互作用データを抽出する際に作成している

**KEGG MEDICUS 医薬品相互作用チェック** [利用法](#) [Top](#)

医薬品リスト

相互作用の有無を確認したい医薬品を追加してください。

商品名または一般名   [医薬品リストをアップロード](#)

**KEGG MEDICUS Drug Interaction Checker** [Help](#) [Top](#)

Drug List

Drug interactions in the Drug List are checked. Please add a drug to the list.

Drug name   [Upload Drug List](#)

クマリン系抗凝固薬

クマリン系抗凝血剤

クマリン系抗凝血薬剤

クマリン系抗血液凝固剤

クマリン系抗血液凝固薬

クマリン系薬剤

ビタミンK拮抗薬

クマリン系抗凝血薬

クマリン系抗凝固剤



B01AA クマリン系抗凝固薬



DGxxxxx

D00564 D01280 D03798

D05457 D07064 D07131

D07135 D07136 D07137

D07969 D08354 D08682

# 薬効を考慮した構造の階層



DGROUP: DG01439

エントリ	DG01439	Structure DGroup
名称	アラビノフラノシル系抗腫瘍薬	
メンバー	DG00685 フルダラビン DG00686 シタラビン DG00688 ゲムシタピン D01223 カペシタピン (JAN) <JP/US> D01633 エノシタピン (JAN) <JP> D01651 塩酸アンシタピン (JAN) D03021 アザシチジン (JAN) <JP/US> D03546 クロファラビン (JAN) <JP/US> D03665 テシタピン <US> D04134 ファザラビン D04233 フルロシタピン D05134 ネララビン (JAN) <JP/US> D06100 テザシタピン D06221 リン酸トリシリピン D06255 トロキサシタピン D09712 エラシタラビン D09722 サバシタピン » expand	
コード	Drug structure map: Antineoplastic - antimetabolic agents [PATH:map07041]	
LinkDB	<a href="#">All DBs</a>	



DGROUP: DG01440

エントリ	DG01440	Structure DGroup
名称	アラビノフラノシル系抗ウイルス薬	
メンバー	DG00645 ビダラビン D00412 ザルシタピン (JAN) D00423 リバビリン (JAN) <JP/US> D01199 エムトリシタピン (JAN) <JP/US> D03981 エルブシタピン D04180 フィアシタピン D06196 トルシタピン D06278 バルトルシタピン二塩酸塩 D06580 デキセルブシタピン D06651 タリバビリン塩酸塩 D07200 イバシタピン D09028 塩酸パロピシタピン D10477 メリシタピン » expand	
コード	Drug structure map: Antiviral agents [PATH:map07044]	
LinkDB	<a href="#">All DBs</a>	



DGROUP: DG00688

エントリ	DG00688	Structure DGroup
名称	ゲムシタピン	
メンバー	D02368 ゲムシタピン D01155 ゲムシタピン塩酸塩 (JAN) <JP/US> D10222 ゲムシタピンエライダート	
コード	ATC code: L01BC05	
コメント	抗腫瘍薬, 代謝拮抗薬 標的分子: ribonucleotide reductase M1 inhibitor [HSA:6]	
LinkDB	<a href="#">All DBs</a>	

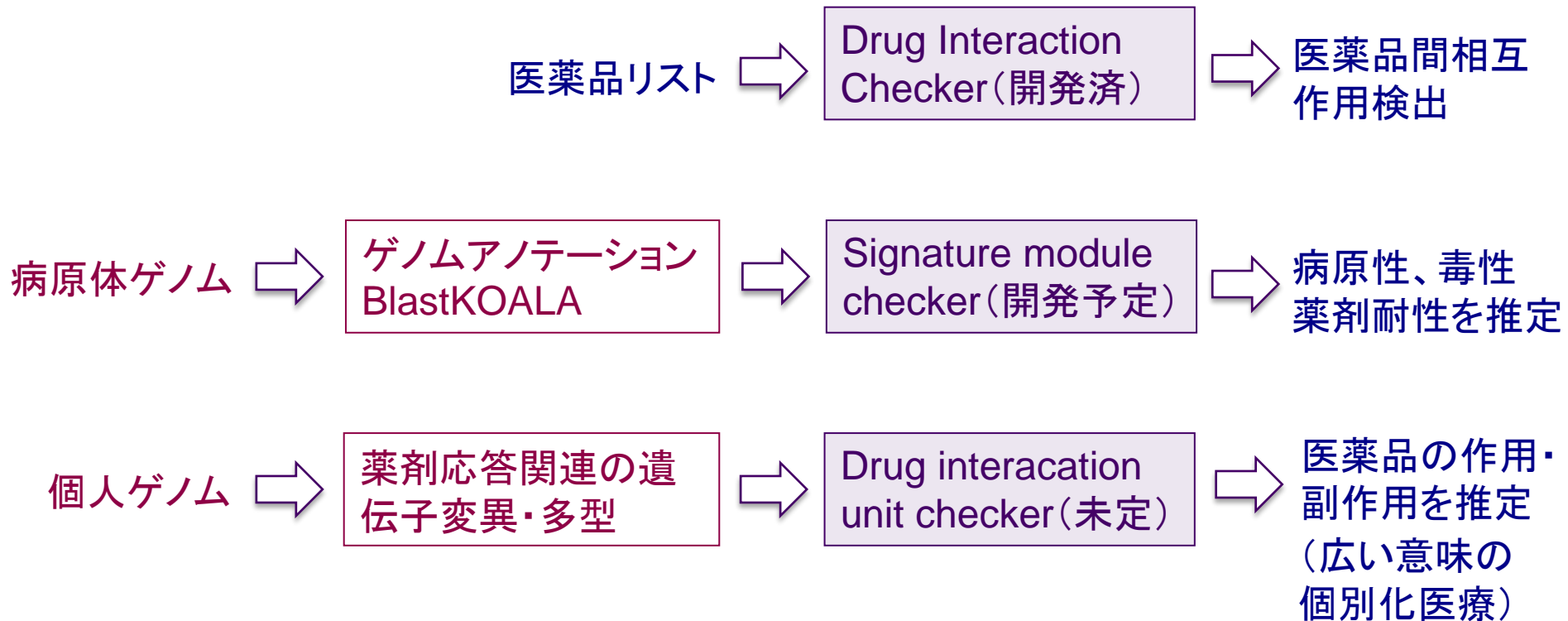


DGROUP: DG00686

エントリ	DG00686	Structure DGroup
名称	シタラビン	
メンバー	D00168 シタラビン (JP16) <JP/US> D03046 シタラビンオクホスファート水和物 (JAN) <JP> D03637 塩酸シタラビン	
コード	ATC code: L01BC01	
コメント	抗腫瘍薬, 代謝拮抗薬 標的分子: DNA polymerase inhibitor [EC:2.7.7.7]	
LinkDB	<a href="#">All DBs</a>	

# 統合インターフェース

KEGG MEDICUS を利用して、ゲノムの情報から疾患・医薬品に関する知見を得るためのツール開発



多型情報の網羅的収集は本計画では行わない

Clinvar など他の DB を用いて Drug interaction unit checker が実現可能かを検討予定

# 機能情報のデータベース化

実験的に得られた遺伝子・タンパク質の機能情報はどの程度データベース化されているか



## KEGG/GenomeNet Feedback Form



Please use the form below to send your question or comment to KEGG and GenomeNet. Please enter information in all fields. (Japanese characters accepted)

Name:

Email address:

(a copy will be sent to you as well)

Affiliation:

Category:

- KEGG (all databases)  
 KEGG tools (BlastKOALA, KEGG Mapper, KEGG API, etc.)  
 GenomeNet tools (BLAST/FASTA, MOTIF, CLUSTALW, KAAS, SIMCOMP, etc.)  
 Copyright permission  KEGG license  Other

Gene function annotation

Organism:

Gene name:

KEGG ID:

PubMed ID:

Comment:

KO の数	18,131
うち PMID があるもの	13,208
うち配列リンクもあるもの	6,884

SwissProt で KO がある割合	69%
SwissProt 中の KO の数	15,159
うち Function のキーワード がついた PMID があるもの	9,697

実験的に得られた遺伝子・タンパク質の  
機能情報を教えてください