

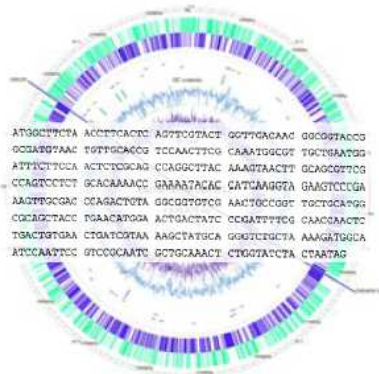
統合化推進プログラム
統合データ解析トライアル
研究成果報告会

マルチオミクスデータを用いた
ゲノム規模代謝モデリングのための
ネットワーク解析システムの開発

理化学研究所 生命システム研究センター(QBiC)
西田孝三

背景: 生命システム理解のための手法について

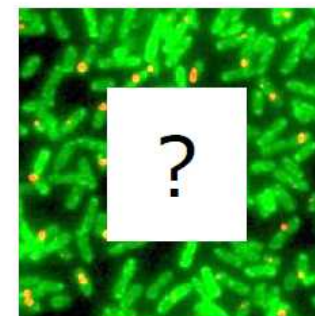
ゲノム



予測

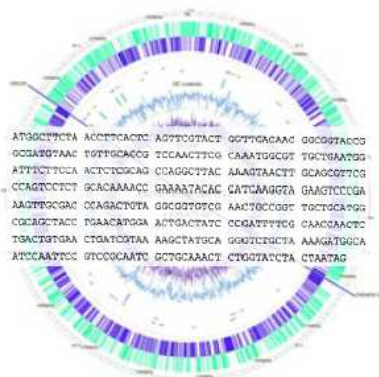
理想

表現型



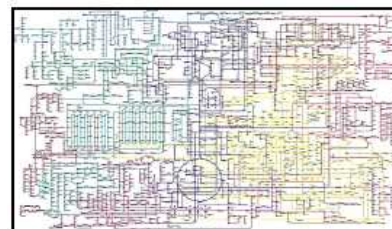
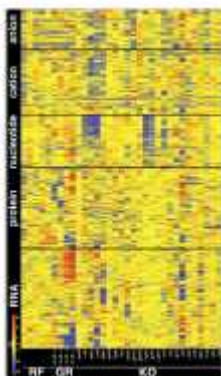
現実

マルチオミクスデータに基づく予測



+

+



基準とその比較対象となる実験条件(遺伝子改変、**薬剤**投与など)を持つマルチオミクスデータ

KEGGに代表されるパスウェイデータベース

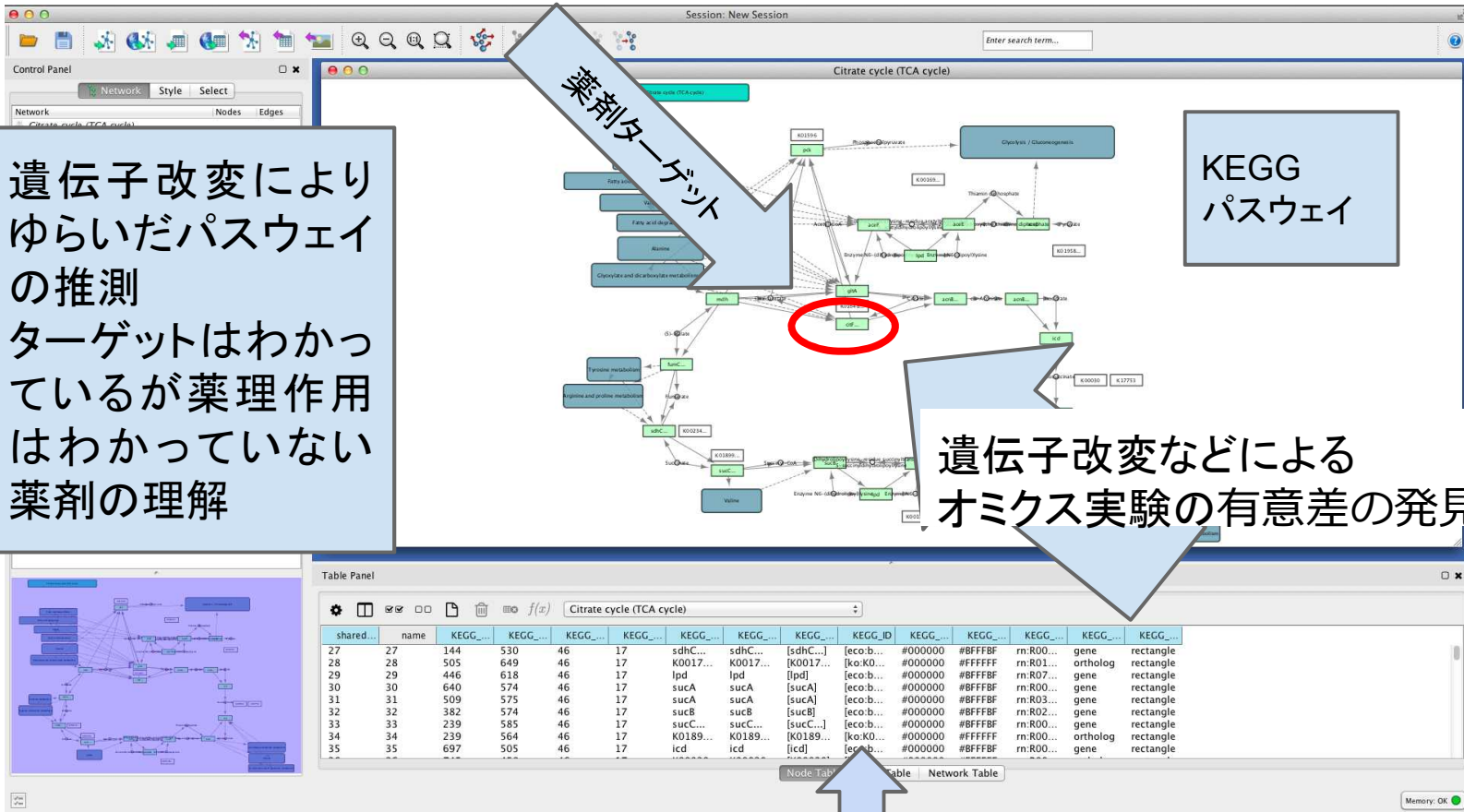
基準に対する違いとして現れる表現型

背景:マルチオミクスデータを用いたゲノム規模のネットワーク解析における障壁

- 下記すべてをKEGGパスウェイにマップしたい
 - KEGG以外のゲノムアノテーション情報
 - KEGG MEDICUS以外のDrug情報
 - 発現もしくはmassプロファイル
- 上記すべてをマップしたパスウェイを可視化、探索したい

これらの要求を満たす汎用的なソフトウェア環境は未成熟

目的: 代謝モデル 薬剤 オミクスパスウェイの統合を容易にするソフトウェア環境の開発



- 遺伝子改変によりゆらいだパスウェイの推測
- ターゲットはわかっているが薬理作用はわかっていない薬剤の理解

ユーザ手持ちのデータの統合

- KEGG KO, MEDICUSに無い遺伝子アノテーション、薬剤情報
- オミクス実験データ

資料: 統合の対象とした情報

- 大腸菌(K-12 MG1655)の代謝モデル iAF1260
- 大腸菌のタンパクをターゲットとするDrugbank中の薬剤情報
- UCSD Palssonグループが公開するマイクロアレイ情報
- 大腸菌(KEGG ID eco)のKEGGパスウェイ情報

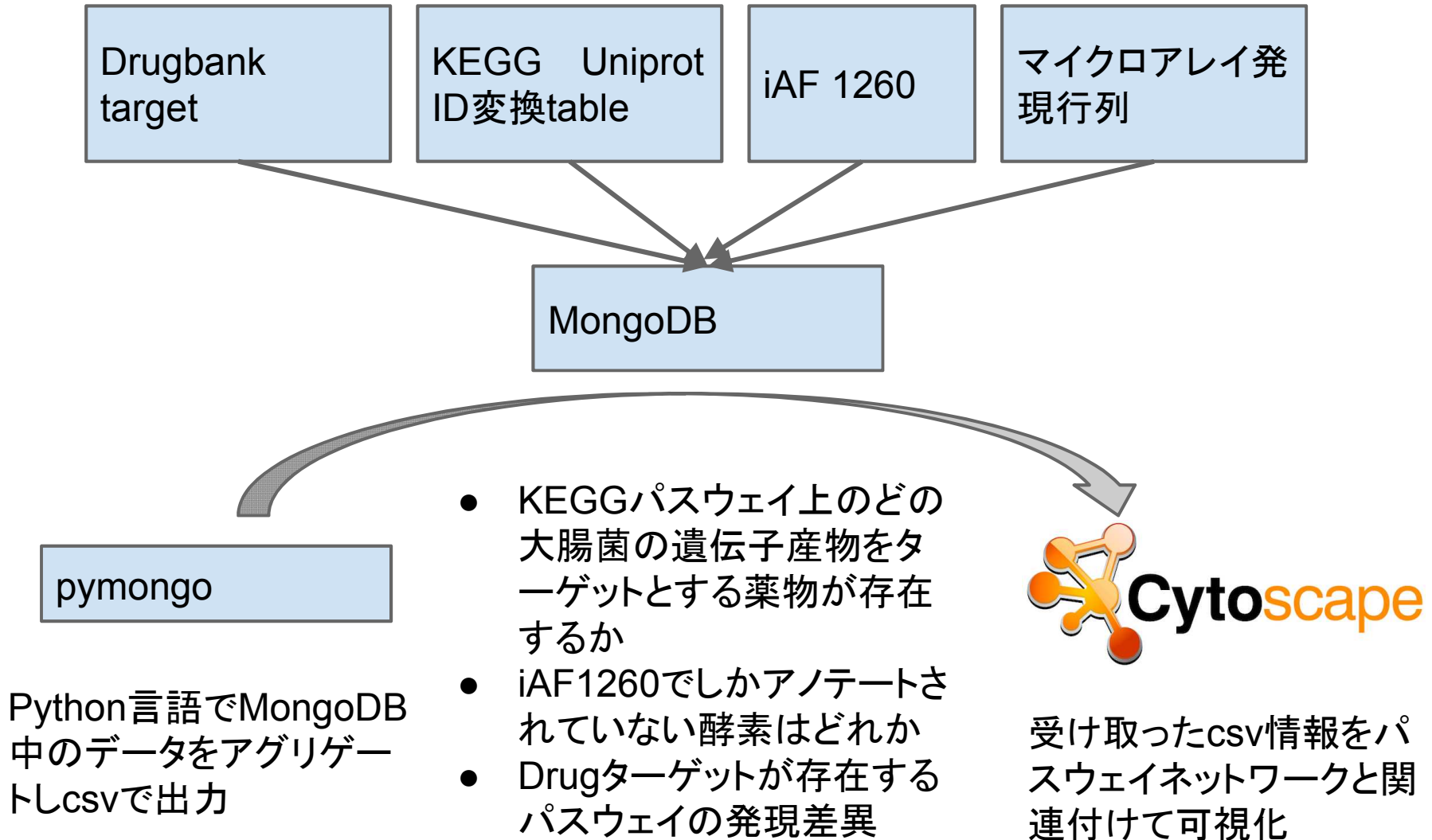
資料: 大腸菌代謝モデルiAF1260

- KEGG(KO)とは異なる手法によって構築された代謝モデル
 - FBAを用いた代謝反応のGap filling
 - 必須遺伝子情報
 - 各反応に対する文献情報
- KEGG(KO)では代謝パスウェイにマップされない48反応を有する
 - KO(eci)による代謝反応数は1153
 - iAF1260でKEGGと対応の取れる代謝反応数は765

資料: Drugbank

- Drug と Drugターゲット(タンパク)のデータベース
- (ecoを含むKOをtargetとする)KEGG Drugには抗生物質(細胞壁合成阻害など)しか無かったため薬剤情報の追加統合対象とした
- version3の全target 4141個中、大腸菌をtargetとするものは360個 (ヒトの2106個に次ぐ)

手法: データ統合のしくみ



研究開発の成果: KEGGscape

The screenshot shows the Cytoscape App Store interface. At the top, there is a navigation bar with the Cytoscape logo, 'Cytoscape App Store', a 'Submit an App' button, a search bar, and a 'Sign In' link. Below the navigation bar, there is a notification banner that says 'Want an easier way to install apps? Click here to learn how!'. The main content area features the KEGGscape app card. The card includes the app icon, the title 'KEGGscape', and the subtitle 'File reader and pathway visualizer for KEGG XML (KGML) files'. Below the title, there are five stars, a '(0)' rating, '224 downloads', and a '3.0' version badge. The card has two tabs: 'Details' and 'Release History'. Below the tabs, there are category tags: 'data integration', 'data visualization', 'local data import', 'pathway database', and 'visualization'. A screenshot of the app's interface is shown at the bottom left of the card. To the right of the app card, there is a 'Download' button with a download icon, followed by the text 'CYTOSCAPE 3', 'Version 0.5.1', 'License Click here', 'Released 12 Feb 2014', 'Works with Cytoscape 3.1', and 'Download Stats Click here'.

一般公開URL <http://apps.cytoscape.org/apps/keggscope>

ソースコード公開URL <https://github.com/idekerlab/KEGGscape>

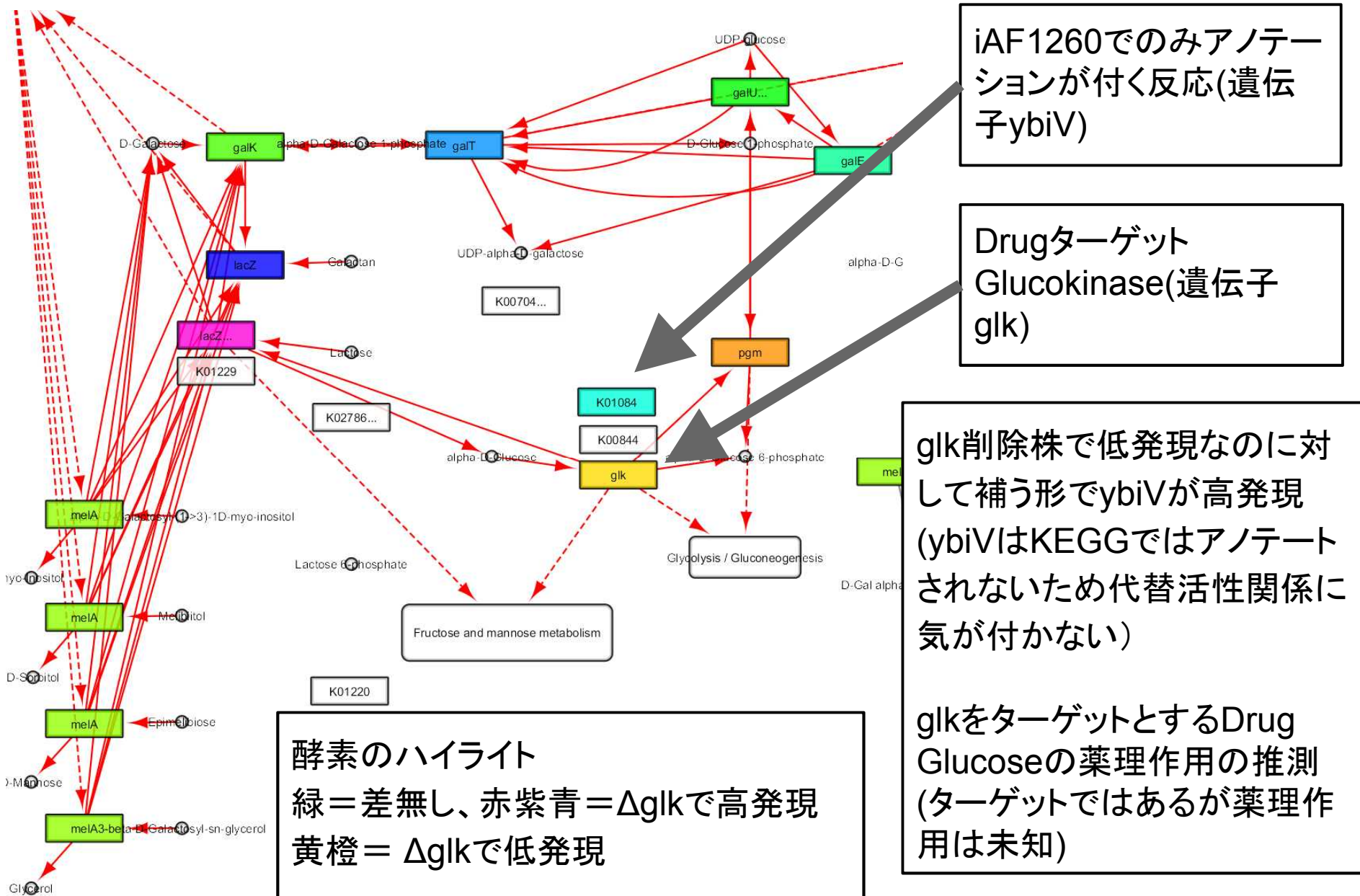
ドキュメントURL <http://keggscope.readthedocs.org/en/latest/>

KEGGパスイエイのXML表現(KGML)をCytoscape(分子間ネットワーク解析プラットフォーム)にインポートする機能を追加する機能拡張

Cytoscapeはメタデータテーブルをネットワークと関連付け可視化を行う

メタデータテーブル中にKEGGのIDを含む任意の情報をパスイエイにマップすることが可能となる

得られる知見の例(Galactose metabolism)



iAF1260でのみアノテーションが付く反応(遺伝子 *ybiV*)

Drugターゲット
Glucokinase(遺伝子 *glk*)

*glk*削除株で低発現なのに対して補う形で*ybiV*が高発現 (*ybiV*はKEGGではアノテートされないため代替活性関係に気が付かない)

*glk*をターゲットとするDrug Glucoseの薬理作用の推測 (ターゲットではあるが薬理作用は未知)

酵素のハイライト
緑 = 差無し、赤紫青 = Δglk で高発現
黄橙 = Δglk で低発現

まとめ

- 代謝モデル(KO以外の遺伝子アノテーション)、オミクスデータを (KEGG MEDICUSが掲げる「ゆらぎ物質」としての)薬剤情報と共にKEGGパスウェイに統合するソフトウェア環境を開発した
- ソフトウェア活用の例として大腸菌の代謝モデルとマイクロアレイデータ、これをターゲットとするDrugのKEGGパスウェイへの統合を行う方法を公開した

今後の課題1

- 一般利用者へのNetwork databaseの公開
 - 本研究では統合例の再現を確約したが、統合後のデータも利用できるようなっていることが望ましい
 - 各データの2次利用公開の可否
 - 最適なデータ形式、データベースの選択(NDEx, Neo4j, MongoDBとの相互補完)
- 各データのLinked open data化
 - 2次利用公開の可否
 - ontology情報の付加
 - SPARQL client

今後の課題2

- iAF1260からiJO1366へ、また他生物種への適用
 - 代謝モデルは大腸菌に限らない
- ChEMBLの追加support、KEGG MEDICUSとの情報比較機能
- KEGGscapeへのPython script担当機能の取り込み

謝辞(敬称略)

- 高橋恒一、海津一成(理研QBiC)
- 大野圭一郎(UCSD)
- 金谷重彦(NAIST)
- 福島敦史(理研CSRS)