

ライフサイエンス知識の階層化・統合化事業
(医薬品・化合物を中心としたデータベースと関連解析ツールの開発・整備を行う)
京都大学

1. 委託事業の9月末時点の判断基準になる目標 (以前にお出しいただいたもの)

- ・ターゲット化合物を全対全で類似構造検索するためのインタフェースを開発。
- ・類似構造検索パッケージの公開
- ・化学反応ネットワーク予測システムの改良
- ・LinkDB 検索インタフェースの改良

2. 9月末時点の達成状況

- ・化合物データセット間の構造比較を全体全で行うためのインタフェース SIMCOMP2 を公開した。
- ・類似構造検索ツール群をパッケージ化し GenomeNet の FTP サイトに公開した (10月)。
- ・化学反応ネットワーク予測システムを改良し、類似反応経路検索システムを開発した (10月)。
- ・医薬品データベースの構造検索結果でアライメント表示をするように改良した。
- ・医薬品データベースのお薬手帖で併用注意・併用禁忌の情報を表示するように改良した。

3. 上記達成状況を踏まえたプロジェクト終了までの目標

- ・化学構造検索、化学反応ネットワーク予測ツールともに、要素技術とそのインタフェースの開発は完了したので、化学情報解析ツールをまとめたインタフェースを作成するとともに、改良点を洗い出す。
- ・統合データベース開発・運用では、開発してきた医薬品データベースの関連情報の整備を引き続き進めるとともに、各種データベースの更新は随時進める。

4. 成果の概要

知識処理技術開発では、化合物データセット間の構造比較を全体全で行うためのインタフェース SIMCOMP2 を開発し、<http://www.genome.jp/tools/simcomp2/>から8月に公開した。また、医薬品データベースに対する化学構造検索システムを6月に公開し、9月には検索結果で構造アライメントを表示するように改良したものを公開した。これらの構造検索で用いているプログラム群をパッケージ化し、GenomeNet の FTP サイト (<ftp://ftp.genome.jp/tools/>) で10月に公開した。反応ネットワーク予測システム PathPred を開発し1月に公開し、反応の基質と生成物のペアを抽出したデータベース RPAIR とペアから抽出した反応パターンのデータベース RCLASS (2010年10月29日版でそれぞれ12,333と2,302 エントリ) を整備している。これらのデータベースを活用した類似反応経路検索システムを開発し9月末時点でテスト公開している。化合物解析システム関係のツールを集めたサイトのアクセス数は、2010年9月で3,016 ノードから約25万アクセス (ロボット含む) であった。

統合データベース開発・運用では、主に医薬品データベースの改良を進めてきた。6月には個人用のおくすり手帖を公開するとともに、各有効成分からターゲットや代謝酵素の情報に関して KEGG の PATHWAY などにリンクを張った。7月には J-GLOBAL へのリンクを追加し、9月にはおくすり手帖で併用注意・併用禁忌の情報を表示するようにした。2010年9月版で医療用医薬品12,239件、一般用医薬品12,042件登録されている。医薬品データベースへのアクセス数は2010年9月で18,380 ノードから約88万アクセス (ロボット含む) であった。