

IDの対応関係に基づく生命情報データベースの統合化 ～リンク自動管理システムとID一括変換システム～

生命情報に関するデータベースはその種類もデータ量も急速に増え続けており、これを有効に活用するための「統合化」が求められている。そこで、データ識別子(ID)の対応関係に着目することにより、効率的かつ持続可能なデータベース統合化を実現する方策を考案した。その方策とは、まず、統合化のカギとなるID(共通IDと呼ぶ)を選定する。次に、共通IDと個別のデータベースに固有のIDの対応関係表を作成し、毎日自動で更新する。そして、IDの対応関係表を利用した検索ナビゲーション機能を実現すれば、実質的な「統合化」が完了する。以上の機能をヒトの分子情報を扱う世界の14機関、19種類のデータベースに対して実現し、リンク自動管理システムおよびID一括変換システム(<http://biodb.jp/>)として公開してその有効性を実証した。この方式で世界中に散在する生命情報関係の主要なデータベースをバーチャルに「統合化」してゆくことにより、生命情報データベースの体系化が実現可能である。

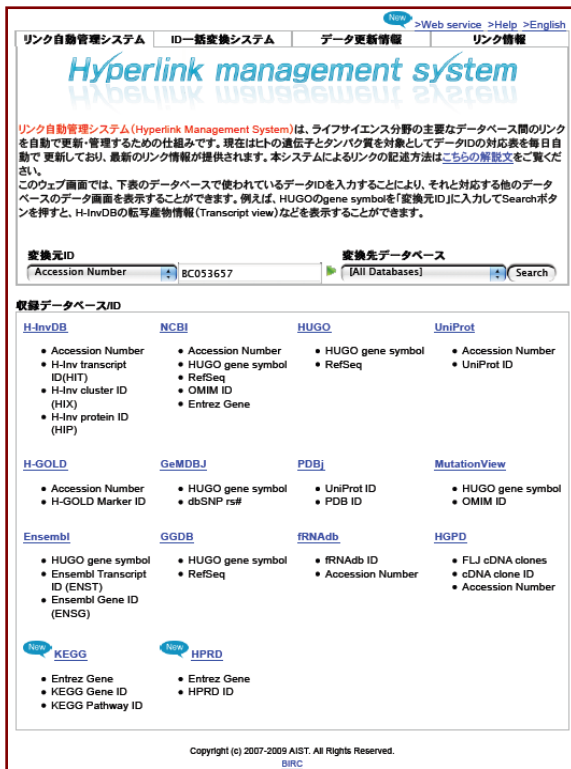


図1 (左). リンク自動管理システム (<http://biodb.jp/>) のトップページ。[All Databases]を選んで実行すると、19種類のデータベースへのリンクを一度に表示できる。

図2 (右). ID一括変換システム (<http://biodb.jp/>)。ヒトの分子情報関連の20種類のIDを一括変換するサービスを提供する。

参考文献

Imanishi T and Nakaoka H (2009) Hyperlink Management System and ID Converter System: enabling maintenance-free hyperlinks among major biological databases. *Nucleic Acids Research* 37 (Web Server Issue): W17-W22 (gkp355).