

パネル5: タンパク質-汎用医薬品相互作用データベースの 開発と知識ベースの構築

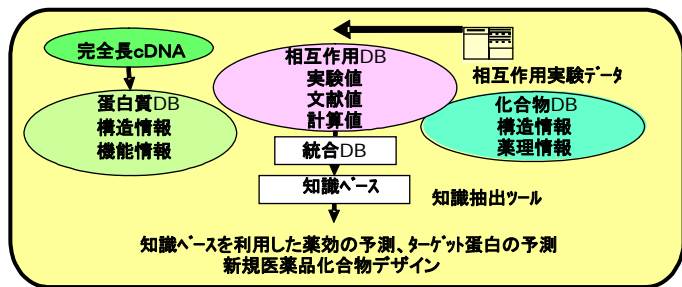


図1 タンパク質-汎用医薬品相互作用
統合化データベースと知識ベース

新しい創薬ターゲットタンパク質を多数発見することを目指して、ヒト完全長cDNAから取得したタンパク質と既存汎用医薬品間の相互作用を網羅的に測定し、得られた相互作用データをタンパク質や医薬品のアノテーション情報と共にデータベース化している。さらに、これらのデータから知識を抽出するためのデータマイニングシステムを構築し、新たな相互作用や薬効の予測に活用可能な知識ベースの構築を行った(図1)。

本委託研究の大きな特徴の一つは、タンパク質と医薬品化合物間相互作用のマトリクスとしての情報が大規模に取得できる点にある。相互作用マトリクスのクラスタリングを行い、得られたクラスターに共通の属性を探っていくことで、各クラスターにとって重要な性質を抽出していくことが可能になる。それらを知識として収集していくことによって、知識ベースが構築できると考えられる。そこで、このような相互作用情報の知識ベース化のフローをインタラクティブに行うことが可能なシステム(Rapid-DD, Rational Analysis of Proteins-Interactions with Drugs for Drug-Discovery)の開発を行った(図2)。

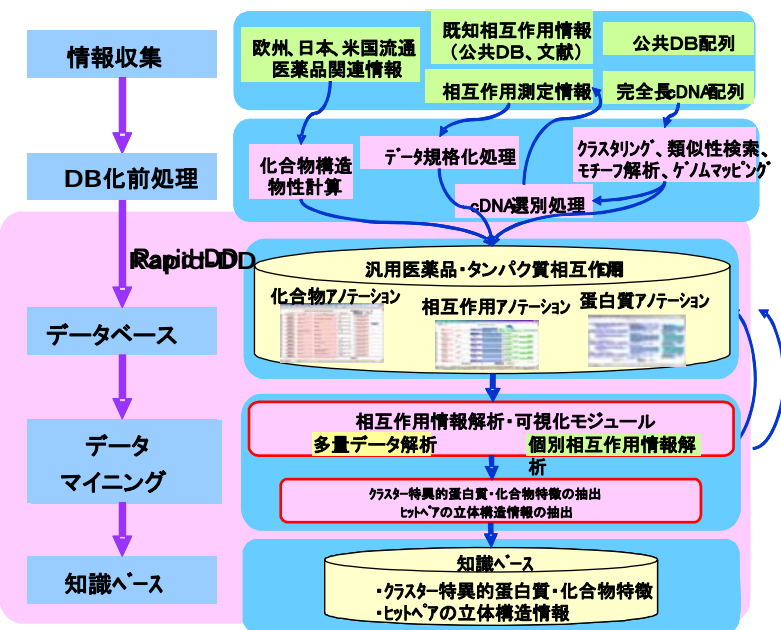


図2 情報収集から1次解析処理・DB化・
二次解析処理・知識ベース化の流れ

相互作用マトリクスからの知識ベース構築フローを図3に示す。相互作用マトリクスのクラスタリングを行い、得られたクラスターとヒットペアに対して、判別分析、ファーマコフォア抽出、薬効や副作用情報との相関解析等を行った。ヒットペアに対しては、モデリングとドッキングを行った。解析結果は知識ベースとして格納した。これらの解析を高精度に行うツールの開発も行った。図4に、HERG蛋白の阻害に対する高精度判別モデルの構築を行った結果を示す。従来法の85%に対して95%の予測精度を実現した。

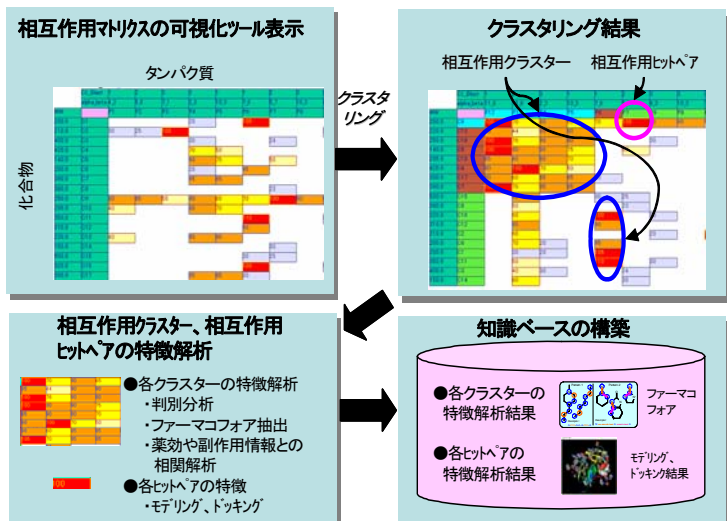


図3 相互作用マトリクスからの知識ベース構築フロー

背景: HERG蛋白は、K⁺イオンチャネルであり、その阻害は、副作用として不整脈を引き起こす。

目的: 医薬品化合物のHERG蛋白阻害の高精度予測
方法: 記述子セットとして、57個の2D記述子と51個の部分構造記述子を用い、サポートベクターマシンを判別モデルとして用いて予測

結果: 判別成功率90~95%を達成(従来法では85%)
結果を論文として、Bioorg Med Chem Lett. 2005 Jun 2;15(11):2886-90に発表

		Separation at pIC50=4.4	Separation at pIC50=6.0
This work	Active	56 / 58 (97%)	24 / 28 (86%)
	Inactive	13 / 15 (87%)	42 / 45 (93%)
	Overall	69 / 73 (95%)	66 / 73 (90%)
Reference	Active	60 / 85 (71%) ^a	5 / 6 (83%) ^b
	Inactive	280 / 329 (85%) ^a	6 / 7 (87%) ^b
	Overall	340 / 414 (82%) ^a	11 / 13 (85%) ^b

^a: Aronov, A.M., Goldman, B.B., *Bioorg. Med. Chem.* 12, 2307, (2004).
^b: Keseru, G.M., *Bioorg. Med. Chem. Lett.*, 13, 2773, (2003).

図4 HERG阻害の高精度判別モデルの構築