

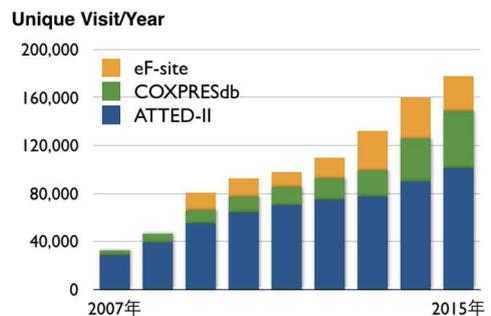
## 平成 27 年度 全体研究開発報告書

1. 補助事業名：創薬等ライフサイエンス研究支援基盤事業  
(創薬等支援技術基盤プラットフォーム事業)
2. 補助事業課題名：構造生命科学データクラウドの構築運用と高度化  
(構造生命科学データクラウドにおけるタンパク質機能推定法の開発と支援)
3. 研究開発代表者：東北大学大学院情報科学研究科 教授 木下 賢吾
4. 研究開発の成果

本事業においては、支援と高度化として以下の目標を立て計画推進を行った。

支援として代表機関（遺伝研）の開発するゲノムや発現量データベースと連携しつつ、P3KとTPRPの成果の利用支援による利用推進を行う。特に、ウェット系の研究者が自ら様々なデータベースを渡り歩けるように、我々が開発を行ってきたDB等も含めて、実践的な講習会を蛋白研・遺伝研と共同して行っていく。また、分担機関独自の支援として、これまで開発を行ってきた独自の手法を利用した構造予測や低分子結合部位の予測を展開していくと共に、遺伝子の共発現データベース等を利用した相互作用パートナー予測などのコンサルテーションを行う。高度化としては、遺伝子共発現ネットワーク、タンパク質間相互作用ネットワーク及び代謝ネットワークを統合し、そこに柔軟な検索機能を付与したデータベースの開発を行う。例えば「ある低分子に着目して、その低分子の代謝ネットワーク上の位置を調べ、その低分子が結合するタンパク質を検索し、そのタンパク質と結合する別のタンパク質を探し、複合体の構造とその低分子の結合部位の関係」を調べたい時に、現在の枠組みでは、KEGGに注目している低分子があってもPDBに対応する低分子がなければ、そのタンパク質と結合するタンパク質を探す段階で「リンク切れ」となり検索ができない。そこで低分子とタンパク質の複合体予測を行う事でリンクを補う必要がある。このような「リンク切れ」は様々な段階で起こっており、本計画ではこれらを一つ一つ補う事で柔軟かつ漏れの無い検索を可能とするデータベースの構築と検索手法の開発を行う。

以上の目標に対して、以下の成果を得ることができた。情報拠点全体の支援としては、実践的な講習会が重要な達成目標であるが、全体の当初目的は十分に達成したと考えている。具体的には、代表機関（遺伝研）を中心として開発を行った VaProS を核として、東北大拠点が高高度化の一環として開発を行った天然リガンドデータベースや遺伝子共発現データベースの実践的な講習会を定期的に行った。ここでは、特に、ウェット系の研究者が自ら様々なデータベースを渡り歩けるように、P3K と TPRP の成果の利用推進として、我々が開発を行ってきたデータベースとして、機能部位の分子表面データベース (eF-site) , 類似性検索による機能部位予測 (eF-seeK) , 遺伝子共発現データベース (COXPRESdb, ATTEDII) 等も含めた幅広い紹介を行った。 分担機関としての個別支援としては、個別の支援成果としては、10 件の個別支援を行い、内 4 件の原著論文の発表を行うことができた。現在も 2 件の論文を準備中である。また、分担機関として、独自データベース COXPRESdb, ATTEDII, eF-site の稼働維持と継続的なコンサルテーションを行い、年間計 18 万ユニーク訪問 (ページビューはデータベースによるが約 7~10 倍) を全世界から得る日本発の存在感の高いデータベースとして利用されるまでになっている (右図)。



高度化では、当初計画の目標とした代謝ネットワーク (KEGG データベースを利用) に関わるタンパク質とそれに結合する基質又は生成物 (低分子化合物) について、それらの複合体構造が不明のものも含めて対象とした網羅的なモデリングを実施し、「リンク切れ」の問題をある程度解決することができた。結

果は、既知の複合体構造を含めてデータベースとして公開を開始した (Natural Ligand Database; NLDB)。また、計画の途中から情報拠点が一体となって、ヒトゲノムの中に見られる多型の解釈に資するシステム開発を行うことになったので、NLDB には、一塩基多型や病気に対応する ID や任意のキーワード等による検索が可能な柔軟な検索機能を実装した。さらに情報拠点全体で開発した構造生物学のデータクラウド VaProS に実装し、タンパク質の相互作用ネットワーク及び遺伝子共発現ネットワーク等の情報を含めた横断的な検索を実現した。また、中間評価でも課題として指摘された点であるが、タンパク質間相互作用に欠けている時間情報 (発現タイミング、遺伝子の共発現情報) の付加に関しても、下図にあるように、当初目的に沿う形で VaProS からシームレスに繋ぐことが出来た。

①要素DBを意識せず検索

②検索結果をチェック

③-1: 詳細画面でNLDBを選択

③-2: 個別の複合体画面で機能部位周辺の変異を確認

④-1: 詳細画面でCOXPRESdbを選択

④-2: 共発現ネットワーク確認

④-3: 代謝ネットワーク確認

Query "hexokinase" hexokinase での検索結果

Gene/Protein result - hits: 8

Gene/Protein	Accession	Start	End	Score	Rank	Gene/Protein	Accession	Start	End	Score	Rank
hexokinase	NP_001161	1	1	100	1	hexokinase	NP_001161	1	1	100	1
hexokinase	NP_001161	1	1	100	2	hexokinase	NP_001161	1	1	100	2
hexokinase	NP_001161	1	1	100	3	hexokinase	NP_001161	1	1	100	3
hexokinase	NP_001161	1	1	100	4	hexokinase	NP_001161	1	1	100	4
hexokinase	NP_001161	1	1	100	5	hexokinase	NP_001161	1	1	100	5
hexokinase	NP_001161	1	1	100	6	hexokinase	NP_001161	1	1	100	6
hexokinase	NP_001161	1	1	100	7	hexokinase	NP_001161	1	1	100	7
hexokinase	NP_001161	1	1	100	8	hexokinase	NP_001161	1	1	100	8

病気の情報とNLDBに入っているもので見ていくと、上位の2つに絞り込み

EntrezGene ID: 3098, Gene Symbol: HK1

Rank	Score	Gene Symbol	Accession
1	10.83	HK1	U03201
2	11.92	HK2	U03202
3	12.37	HK3	U03203
4	17.89	HK4	U03204

③-1: 詳細画面でNLDBを選択

③-2: 個別の複合体画面で機能部位周辺の変異を確認

Residue	AA	UniProtKB	dbSNP	Type	Disease name
N_538	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_539	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_718	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_719	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_720	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_721	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_722	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_723	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_724	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_725	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_726	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_727	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_728	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_729	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_730	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_731	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_732	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_733	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_734	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_735	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_736	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_737	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_738	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_739	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_740	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_741	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_742	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_743	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_744	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_745	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_746	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_747	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_748	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_749	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_750	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_751	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_752	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_753	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_754	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_755	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_756	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_757	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_758	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_759	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_760	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_761	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_762	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_763	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_764	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_765	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_766	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_767	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_768	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_769	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_770	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_771	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_772	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_773	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_774	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_775	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_776	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_777	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_778	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_779	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]
N_780	LEU	SEN			Disease: hexokinase deficiency (HK deficiency) [MIM 235705]

④-1: 詳細画面でCOXPRESdbを選択

④-2: 共発現ネットワーク確認

④-3: 代謝ネットワーク確認

KEGG ID Title | #genes | Link to the KEGG's database

map01100	Metabolic pathways	12	
map01010	Glycolysis / Gluconeogenesis	9	
map01020	Pyruvate and oxaloacetate metabolism	3	
map01030	Amino sugar and nucleotide sugar metabolism	2	
map01040	Starch and sucrose metabolism	2	

Genes directly connected with HK1 on the network

Gene	Cor	Symbol	Function	Accession	Entrez Gene
16.5	0.92	PFKM	phosphofruktose mutase 1 (Bran)	[Entrez]	5223
17.1	0.27	PFKB2	transport and gtp-organization 2 homolog (Chromop)	[Entrez]	128989
28.8	0.85	GLAC1	glucanase domain-containing 1	[Entrez]	10422

以上のように、まだ改良の余地のあるインターフェースではあるが、当初想定以上のデータベースとシステムの開発を行うことができたと考えている。その結果、タンパク質間相互作用に大きな影響を与えながら、濃度の調整で変異の影響を回避し高頻度で観察される面白い変異 (Nishi et al, Pro Sci, 2016) を報告することもでき、様々な低分子を認識することができる特徴的な結合部位 (Sociable Binding Site) を同定 (Murakami et al, Pro Sci, 2013) することができるなど、単なるデータベース構築にとどまらない科学的な知見を得ることもできた。