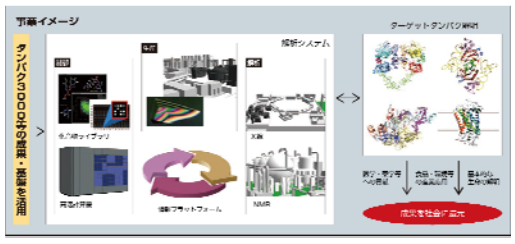
ターゲットタンパク研究プログラムの知的基盤「情報プラットフォーム」

○SUGAWARA Hideaki[1)], NAKAMURA Haruki[2)], KANEKO Akihito[3)], OHNO Yoshie[4)]

1) National Institute of Genetics., 2) Institute for Protein Research, Osaka University.
3) Life Science Research Laboratory, Central Research Laboratory, Hitachi, Ltd.
4) Graduate School of Agricultural and Life Sciences, The University of Tokyo.

Contact us: portal@nig.ac.jp

文部科学省ターゲットタンパク研究プログラムは、基本的生命、医学・薬学、食品・環境の３つの領域におけるターゲットタンパク研究と、タンパクの生産と解析、化合物による制御、そして情報プラットフォーム(情報PF)の４つの技術開発研究における４３課題で構成されています。情報PFは、プログラム内の多様なグループの間での情報共有をもたらすべく設計されました。さらに、ターゲットタンパク研究プログラムの研究成果を一般に広く公開することも目的としています。



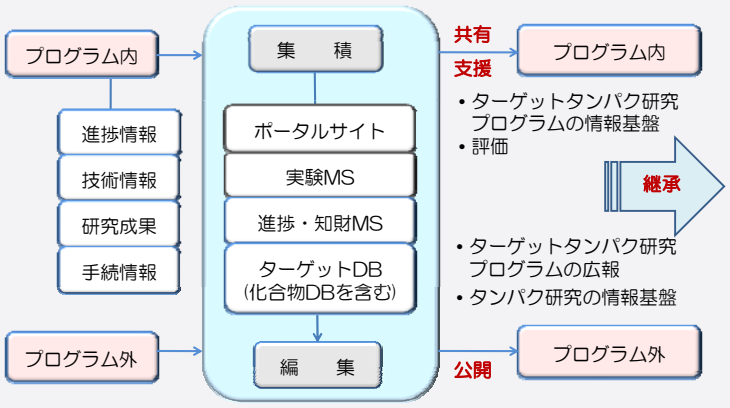
情報PFの役割は２つ

「タンパク3000プロジェクト」と「ターゲットタンパク研究プログラム（TPRP）」の研究成果を広く一般に提供すること

☞ 公開サイト

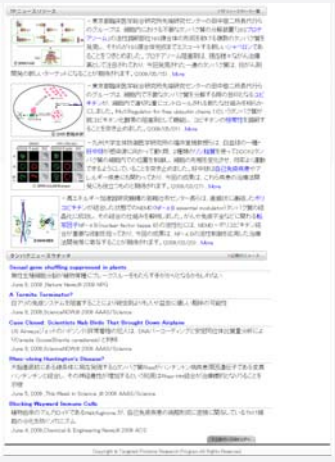
情報を獲得し共有する情報システムを用意してTPRPを推進すること

☞ 共有サイト，情報マネージメントシステム（MS），データベース（DB）



公開サイト

http://www.tanpaku.org

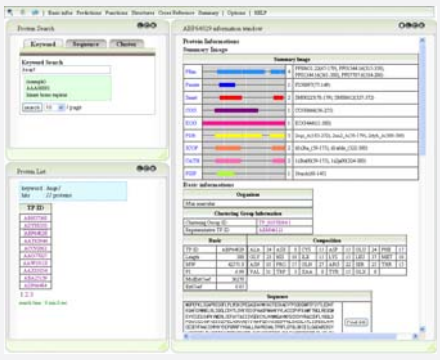


共有サイト

TPRPの各参加者ごとの役割によってアクセス権限が設定されています。ここから、配列DB，構造DB，実験MS，成果登録，TanpakuWiki，ビームタイム予約，TPリンク，進捗登録，評価支援が利用できます。これらのいくつかは近日、公開サイトでも公開致します。

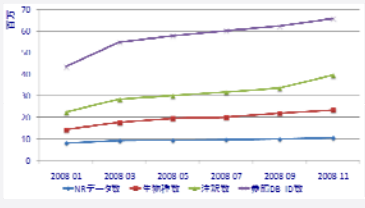
配列DB

http://sqpr-web.genes.nig.ac.jp/UI/top.html



世界の主要なアミノ酸配列DBから毎月データを収集・統合した非冗長データセットを作成し、各種タンパク質情報を付加した配列データベースを運営しています。現在1000万件超の非冗長アミノ酸配列、2,500万件超の生物種情報、6,500万件超の参照DB_IDを収集しています。

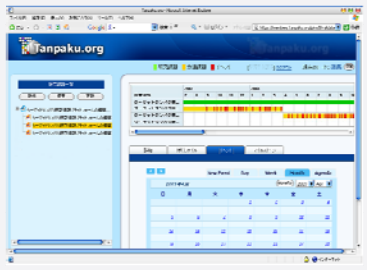
配列DB：各種データ数の推移



成果情報の集積と活用

共有サイトの「進捗登録」または「成果登録」から登録した情報は内部ならびに外部評価の基礎資料となり、また「成果報告書」の半自動生成にも活用されます。
「ビームタイム予約」および調整をWeb上で行うことで、効率的な研究を支援しています。

進捗MS



成果報告書出力

研究成果一覧

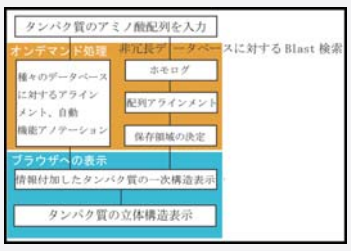
ビームタイム予約

タンパク質解析支援

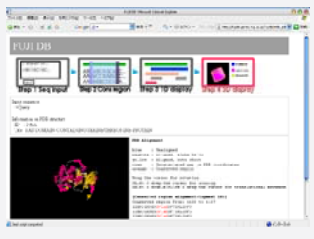
情報PFでは、よく知られている公共データベース群から全アミノ酸配列を獲得して抽出した1000万を超える非冗長配列に情報を付加した「配列DB」、任意のアミノ酸配列に対して配列の特徴分析に加え立体構造や天然変成(disorder)領域の予測も行う「構造DB」、参加機関の成果をDB化してわかりやすく公開する「個別DB」など情報の視点からタンパク質解析を支援しています。

構造DB

http://fujidb.genes.nig.ac.jp/fujidb/

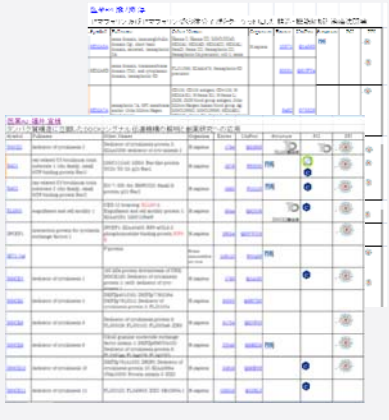


任意のアミノ酸配列に対して機能予測や立体構造の予測を行うオンデマンドシステムです。データベースを頻繁に更新し、計算時間を短くすることにより、アミノ酸配列が明らかになった任意のタンパク質について利用者は迅速に最新の立体構造、立体構造を利用した機能予測などの情報を得ることができます。



TP table集

各ターゲットタンパク研究グループで対象としているタンパク質をリストアップし、基本情報から、現在までに得られた構造情報，スクリーニング情報，タンパク質間相互作用情報，発表論文情報を集めたテーブルを作成しています。グループごとの研究対象を俯瞰しやすくするのを目的としています。

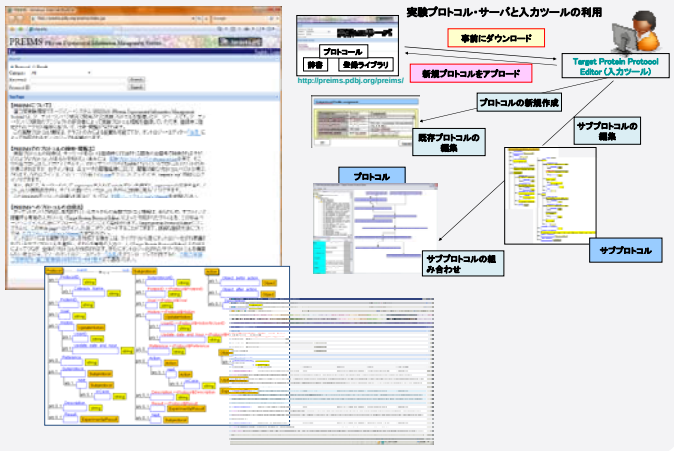


実験情報の定義と登録

http://preims.pdbj.org/preims/index.jsp

蛋白質実験情報マネージメント・システム（PREIMS: PRotein Experimental Information Management System）は、ターゲット・タンパク研究で開発された実験プロトコルを整理したデータベースです。ターゲット・タンパク研究のプロジェクトの研究者によって実験プロトコル情報を登録していただき、登録時に指定されたアクセス権限に基づいて、検索・閲覧がなされます。
（４月時点で63件登録済）

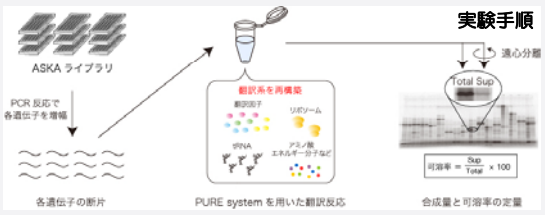
実験MS



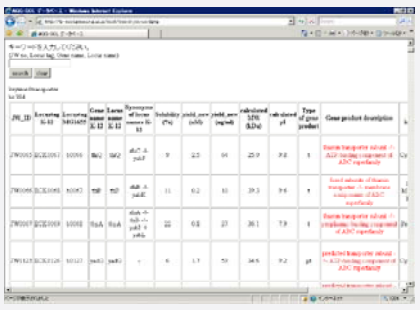
個別DB

e-Sol (Solubility database of all *E.coli* proteins)

再構築型の試験管内タンパク質合成系(PURE system)を用いて大腸菌の全タンパク質を発現させた際の、凝集の度合いと合成量の実験結果をまとめたものです。東京大学大学院新領域創成科学研究科の上田卓也教授らのグループ(技術開発研究生産C)が行った研究成果*をデータベース化しました。



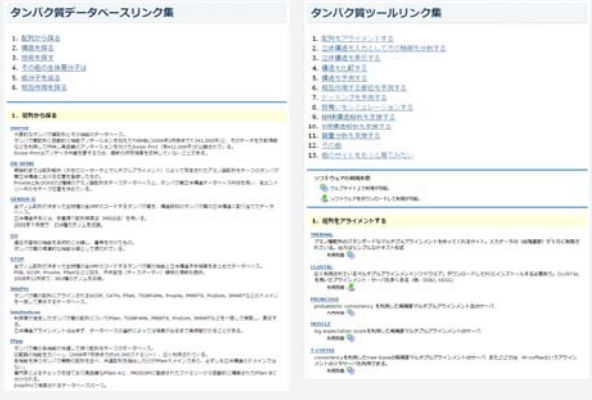
http://tp-ecoli.genes.nig.ac.jp/ecoli/



*Niwa, T. *et al.*, (2009) *PNAS*. **106**, 4201-6

TPリンク集

タンパク質解析に有用なサイトやツールを揃えた独自のリンク集を作成しました。データ解析の手助けになるよう、目的別に分け、それぞれに簡単なコメントを加えました。
（近日中に公開サイトへ移行予定）



遺伝研 菅原研究室の全メンバ