

構造生命科学データクラウドの構築と運用

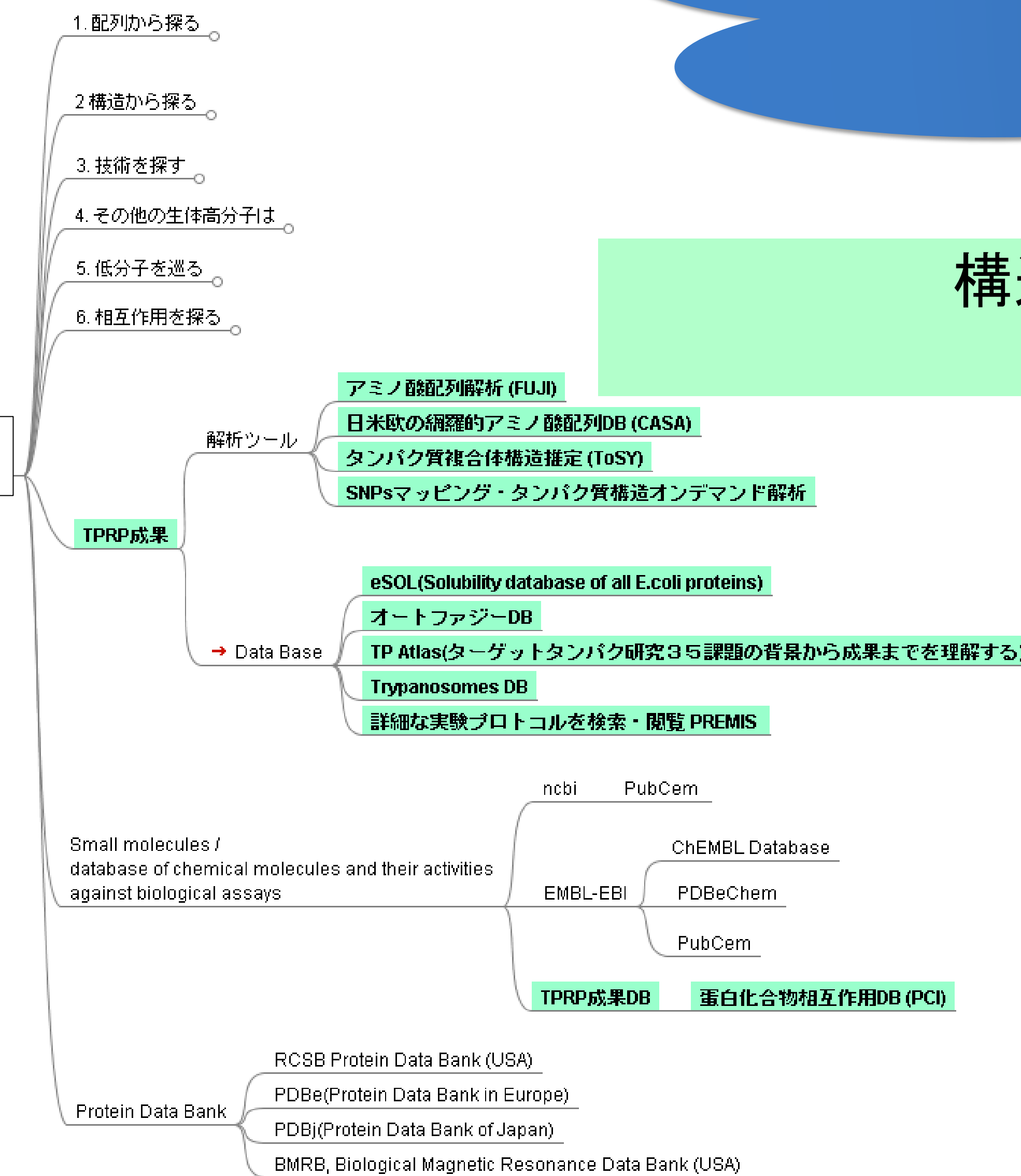
代表機関 国立遺伝学研究所(五條堀 孝)

<http://p4d-info.nig.ac.jp/>

わが国が今まで投資してきた関連分野のプロジェクトの成果を、構造生物学にとどまらず生命科学全体に浸透させ、関連する企業などの波及効果も考慮して、創薬等に向けた「発想を支援する俯瞰機能」と「問題解決を可能とする情報環境」を提供し、わが国における創薬等を支援する情報基盤「構造生命科学データクラウド」を確立する。

構造生命科学データクラウドはインターネットから精選した多様な分野のデータベース・ツール群(左図に概要)と独自に構築する特徴あるデータベース・ツール群(下図に例示)の融合

タンパク質情報



eSOL (Solubility database of all *E.coli* proteins) ターゲットタンパク研究プログラムにおいて、東京大学上田卓也研究室と共同で構築し今後真核生物のタンパク質へと拡張の予定

オートファジーデータベース (オートファジー関連のタンパクとそのホモログの情報を網羅。日々更新される関連論文リストも用意) ターゲットタンパク研究プログラムにおいて国内オートファジー研究グループと共同で開発し、拡充中

個々のタンパク質の結晶化支援ならびに巻き戻し技術の俯瞰を可能にする高度なデータベースを、東大農学部田之倉研究室ならびに阪大蛋白研高木研究室と共同で構築し、タンパク質結晶化の支援に供する

ゲノム情報からタンパク質構造情報を介してフェノタイプとの相関に関する仮説創出・検証の「場」を提供することにより、実験研究を加速

「場」となる解析パイプラインの例 -SNPsとタンパク質の機能変化・疾患の相関-

機能的SNPの絞り込み

例) 特定の疾患を対象
アミノ酸置換型SNPからタンパク質の機能にダメージを与えそうな候補SNPの予測

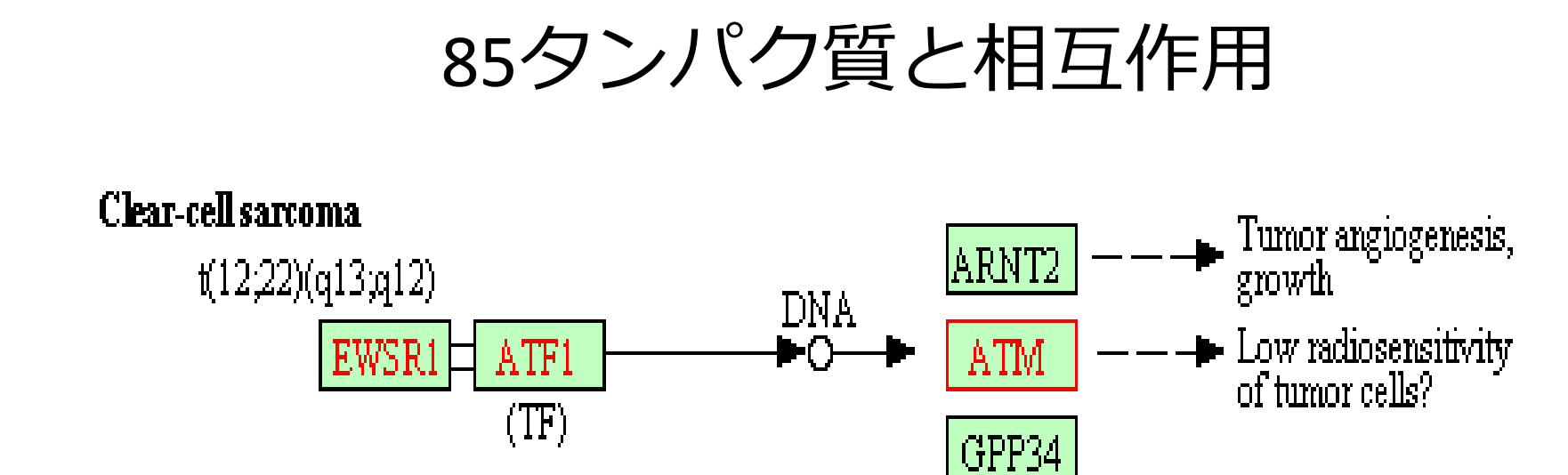
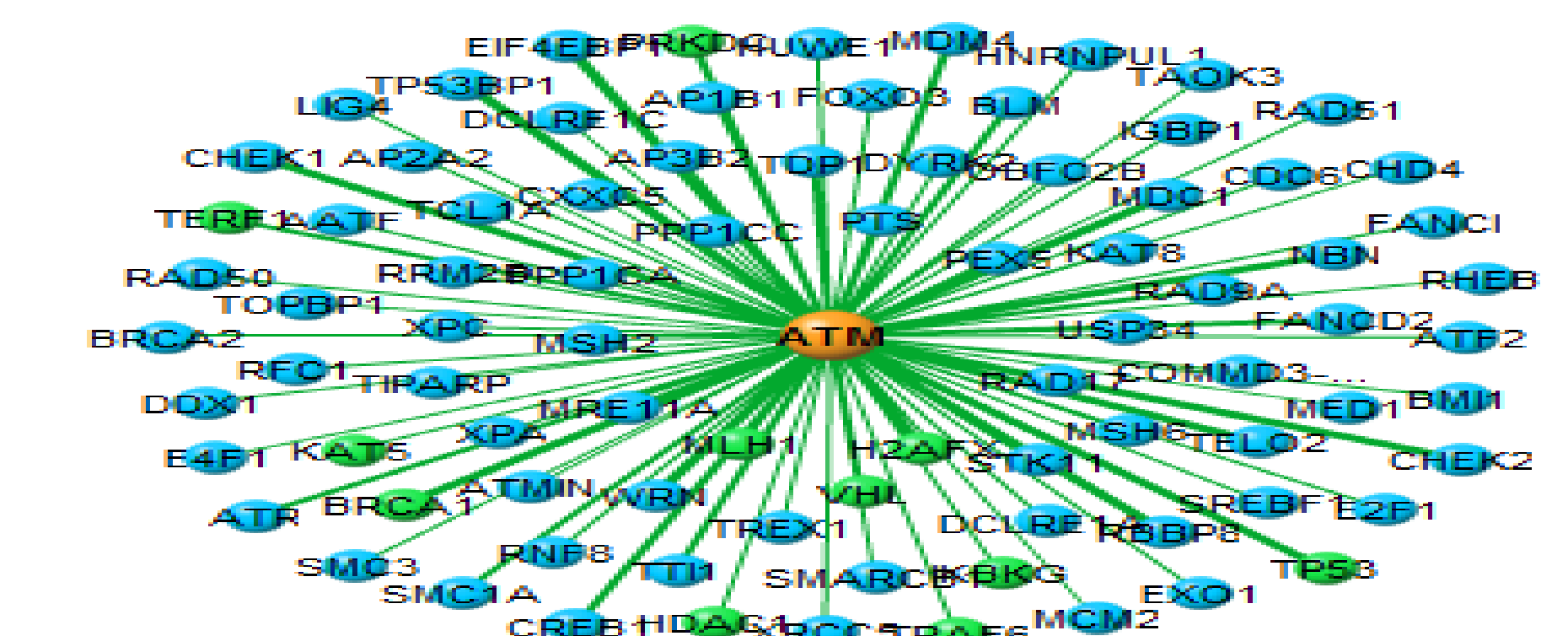
Protein interactionへの影響 (GNP: PPI)
発現情報を用いたタンパク質相互作用の絞り込み

KEGG(代謝パスウェイの相互作用)

e.g. ICGC (RIKEN) データ
Mutation ID: sm130478

日本人、男性、58歳、肝臓がん

ATM
11番染色体 108200958
A->C
Q->P
Non_synonymous_coding
新規



Transcriptional misreguration

ゲノム情報

遺伝子発現情報

構造生命科学データクラウド

研究者
(変異解析、臨床解析)



遺伝研

解析パイプライン
変異 (SNV/SNP) の機能予測
パスウェイ、疾患アノテーション

公共DB

様々な切口から
情報検索

検索インターフェイス

統合アノテーションDB
ゲノム上の変異と構造・機能・疾患への影響を包括的に整備

データ連携

大阪大学