

開発者から

疾患ヒトゲノム変異の生物学的機能注釈を目指した多階層オームクスデータの統合

菅野 純夫

東京医科歯科大学難治疾患研究所 非常勤講師

DBKERO (Database of Kashiwa Encyclopedia for human genome mutations in Regulatory regions and their Omics contexts) はヒトゲノム中に見出された多型、変異における生物学的意義を検証するために、周辺領域の多層オームクスデータを集積したデータベースである。特にその遺伝子発現制御かく乱の可能性についての情報を抽出すべく、広く、エピゲノム、トランスクリプトームのデータを内外から収集、掲載している点に特色がある。また申請者らが近年、実践している新規ゲノム解析技術から得られたデータセットについてもいち早く生データとして公開している(図1)。今回、その最新の更新についてUpdateを報告する。

新規データについて、今年度、我々は、日本人ゲノムについて合計13,368人のゲノム多型/変異のデータを拡充した。これは様々な国内GWAS研究により収集されたものである。これに加えて、参照エピゲノムデータを455 datasets (67人の健常人組織)を公開した。これは、International Human Epigenome Consortium (IHEC)におけるCREST日本チームの貢献としてデータ産生、公開されたものである。新規ゲノム解析技術を用いたものについては、がん細胞を中心としたシングルセル解析の結果について30データセットを公開した。これらは肺腺癌細胞株について、種々の薬剤刺激時での細胞応答多様性を検出すべく収集されたものである。さらに、合成的、物理的ロングリード解析として、Nanopore/10x Genomicsのデータを公開している。データコンテンツの検索例を図2に示す。本データベースは、<http://dbkero.hgc.jp/> から無償で全データを取得可能である。

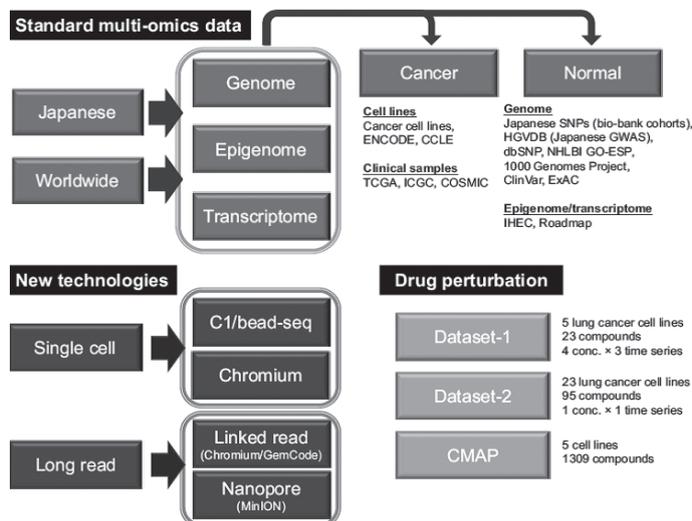


図1 DBKEROのデータコンテンツ

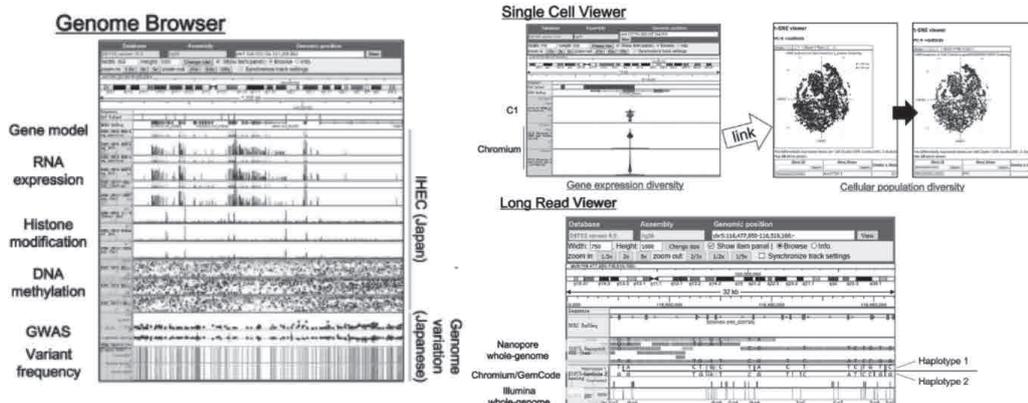


図2 DBKEROのデータ検索例

参考文献

Suzuki A et al. DBTSS/DBKERO for integrated analysis of transcriptional regulation NAR DB issue 2018