

報告番号	※ 第 号
------	-------

主 論 文 の 要 旨

論文題目 創薬研究効率化のための
高次元情報のインフォマティクス

氏 名 河 合 駿

論 文 内 容 の 要 旨

近年、ライフサイエンス研究から得られる情報は爆発的に増大している。ビッグデータとも表現される膨大かつ高次元な情報を高度なコンピュータ解析によって活用しようとする試みは、現在の新薬開発における成功確率の低さを克服する極めて重要な創薬技術として位置づけられつつある。

創薬研究の中で得られる情報には、オミクスデータのように生体内の多種多様な分子を網羅的に計測したデータの他にも、分子レベルでの構造データ、合成経路や製造工程などのプロセスデータ、細胞アッセイから得られるデータ、画像などのハイコンテントデータなどがあり、データの種類と次元は人間が直感的に把握できる範疇を遙かに超えている。近年得られる情報は急速に多様かつ複合的になってきており、生命現象に迫るための解析理論に加え、多様な解析目標に合わせて実践的に研究推進するツールとしての解析手法の開発が求められてきている。

本論文では、創薬研究における化合物スクリーニングという研究目標に焦点を置き、高次元情報の活用によって研究の効率化を促進するための情報解析の基盤的理論及びモデル化技術の構築を目指した。特に高次元情報として、セルベースアッセイ中の細胞画像から得られる表現型情報という新しいタイプの情報について、その実践的な応用及び有効活用について研究を深めた。

本論文は 5 章から構成され、序章として創薬研究における情報解析（バイオインフォマティクスやデータサイエンス）の最新動向、高次元情報を活用した創薬研究の現状と課題、そして近年の創薬スクリーニングのトレンドと求められる技術についてまとめた。

第 2 章、第 3 章では、創薬研究における化合物評価の効率化における細胞形態情報解析の有効性について論じる。具体的には、薬効評価基盤技術の確立、解析手法の確

立、及びその応用の可能性について研究事例をまとめた。

第4章では、情報理論的な観点から、細胞形態情報の特徴と取扱、そしてその有効な解析手法について論じる。具体的には、細胞形態情報を定量的情報として活用するための新規アルゴリズムに関する研究事例をまとめた。

第5章は本研究の締章として、高次元情報のインフォマティクス解析が有する創薬研究効率化における可能性についてその他の研究事例と合わせて考察すると共に、その将来性および課題についてまとめた。

第2章では、細胞形態情報解析の応用により、神経成長因子(Nerve Growth Factor : NGF)の遺伝子発現を誘導する薬剤評価系を効率化した事例を示す。

従来、NGFなどの神経分化を促す候補薬剤分子の評価では、長期間(5日間)のアッセイを薬剤応答細胞と表現型検出細胞の2種類を利用する必要があった。薬剤応答にはアストロサイトーマのような分子の添加に応答してNGF産生細胞が用いられるが、産生タンパク質が微量であることや、NGFの生体内での薬効を評価するためにも、神経突起の伸張の表現型を示す細胞(ラット副腎髄質由来褐色細胞腫(PC-12細胞))に曝露する評価プロセスが必要とされていた。このような複雑かつ長期間のアッセイは、スクリーニング効率を大きく低下させると共に、結果の不安定性という課題を抱えていた。

本研究では、このアッセイ工程中において用いられる両方の細胞について細胞形態情報を定量化・分析することで、NGF誘導能が知られる天然化学分子LyconadinBの性能をいかに効率的かつ定量的に検出できるかを検証した。結果、通常の計測対象となるPC-12細胞ではなく、NGFを産生するアストロサイトーマの形態情報を直接定量化することによって、実験開始後たった9時間で形態特徴量の変化として薬剤機能性が検出できることが見出された。また形態特徴量という高次元情報の変化をコンピュータでモデル化することによって、通常5日間かかるNGF誘導能のアッセイを、評価開始6~9時間という早期に85.7%の精度で予測できることを示し、アッセイ期間を13倍短縮できることを示した。

第3章では、異なる作用機序を持つ抗がん剤候補分子を、非破壊的に得られる細胞形態情報のみから評価・判別する解析手法の開発について示す。

従来、抗がん剤などの薬剤候補分子のスクリーニングは、評価対象である細胞を破壊的に評価するエンドポイントアッセイが主流であり、薬効解析を短時間化しスループットを向上に貢献してきた。しかし、経時的な薬効評価が難しく、微量・長期間曝露における毒性評価やがんの種別に対する応答差などのわずかな違いの評価は難しく、分子スクリーニングの感度や評価能を向上させることに限界があった。また薬効評価用の蛍光分子などはスクリーニングコストとしても大きい課題があった。

本検証では、このような分子スクリーニングの高度化・低コスト化を目指し、培養細胞の画像から得られる高次元表現型情報(細胞形態の経時的变化)を用いて、抗が

ん剤を作用機序ごとにプロファイリングできるかを検証した。具体的には、NCI60 がん細胞パネルの 5 種類のがん細胞（A549, A-498, NCI-H23, SK-MEL-28, PC3）に 9 種類の抗がん剤を投与し、その細胞形態高次元情報プロファイルから作用機序を予測する解析法の確立を目指した。

結果として、主成分分析によって、各薬剤の濃度別条件における細胞形態情報という点を線で繋いだ「軌跡」として、抗がん剤作用機序の違いを可視化できる手法を確立した。即ち、細胞は、抗がん剤の薬剤機序に応じて特徴的な形態変化を示しており、「濃度別」「細胞別」「経時別」に対応した細胞形態変化を全て取得し高次元情報として活用すると、培養細胞の画像情報のみから薬剤の作用機序のタイプを判別できる可能性が示唆された（予測精度 87.5%）。本研究で、細胞形態情報という高次元情報を活用した、経時的非破壊薬剤プロファイリング法を確立することができた。

第 4 章では、細胞形態情報解析で用いる「形態情報」という情報量の意味と有効性についてより深く検証し、解析法の理論的改善を行った。

細胞形態情報は、培養細胞の細胞形態情報を網羅的に数値化したものであり、細胞集団の変化などのわずかな生物変化を数値情報として定量化することが可能である。しかし同時に、膨大な数値情報が経時的に得られるため、情報量として独立性の低い情報を多く含み、次元のみが増大し、様々なノイズを含みやすくなるという欠点が存在した。しかし、細胞形態情報は極めて新しく独自性の高い情報であり、情報理論的観点からこの情報量の取扱の最適化について詳細に検証した例はこれまで無かった。

本検証では、不均一な細胞集団から得られる形態特徴量を構造化された情報として活用するモデル化技術の有効性を検証した。構造化とは高次元情報を部分的にグループ化し、グループとしてまとめた変数群の変化や傾向をモデルに組み込む技術である。結果、従来は部分的かつバラバラな情報として利用していた高次元情報に対して、幾つかの情報をグループ化したモデルは解釈可能かつロバストなモデルとなることが示唆された。

本論文の研究成果より、創薬開発におけるセルベースアッセイにおいて、細胞画像から得られる細胞形態情報という高次元情報をいかにインフォマティクスとして処理・解析すべきか、という理論構築、及び、その薬剤探索・プロファイリングにおける有効性を実証することができた。また本研究からは、様々な高次元情報を活用する「情報工学的解析」と「実験的実証」を融合的に循環させる研究によって、従来の創薬探索・検証が大幅に効率化できる可能性が示唆された。このような情報解析と実験的検証の融合は、創薬研究のみならず、今後のビッグデータの時代におけるライフサイエンス研究の加速と進化において極めて重要な貢献となることが期待される。