

厚生労働行政推進調査事業費補助金（化学物質リスク研究事業）
新型毒性試験法とシステムバイオロジーとの融合による有害性予測体系の構築
（H30-化学-指定-001）
平成 30 年度 分担研究報告書

分担研究課題：システム毒性解析の人工知能化

研究分担者： 北野宏明
特定非営利活動法人システム・バイオロジー研究機構 会長

研究要旨：

システム毒性では、一連の解析手順の高度な連携と同時に、大規模データベースから多くの情報を抽出し、それを解析へと結びつける必要がある。本分担研究では、深層学習（Deep Learning）を用いて膨大な遺伝子変動データから有意に変動した遺伝子を高精度で自動同定させる技術、ならびに解析パイプラインの連動強化を行った。

研究協力者：

長谷武志（システム・バイオロジー研究機構）
Natalia Polouliakh（ソニーCSL）

A. 研究目的

システム・レベルで毒性を理解するには、膨大な実験データを格納したデータベース、文献、数値モデルなどを統合的に解析する必要があり、大規模かつ複雑なデータを意味のある形で解析するには、深層学習（Deep Learning）やテキストマイニングなどを含めた一連の人工知能（AI）アルゴリズム群の連携が有効である。さらに、複数の解析ツールをスムーズに連動させる必要がある。本分担研究では、一連の解析過程の AI 化を実施し、ツール間連動を強化することで、高度な AI 駆動型システム毒性学基盤の構築を推進する。

B. 研究方法

システム・レベルで毒性を理解するには、膨大な実験データを格納したデータベース、文献、数値モデルなどを統合的に解析する必要があり、大規模かつ複雑なデータを意味のある形で解析するには、深層学習やテキストマイニングなどを含めた一連の人工知能（AI）アルゴリズム群の連携が有効である。本分担研究では、一連の解析過程の AI 化を実施する。

具体的には、H30年度は一連の解析工程から AI 化が可能と思われる工程を洗い出し、開発マイルストーンを設定する。また、その一部に関してのプロト

タイプを開発する。また先行研究による SHOE とその関連ツールを強化し、それらを AI 化する方策を策定、設計を行う。

H31 年度は開発マイルストーンに従い、深層学習とテキストマイニングを中核に Percellome データベースの AI 化を進める。また SHOE 及び関連ツールの強化を継続し、AI との連動を可能とする API の定義と実装を行う。

H32 年度は Percellome データベースの AI 化を完了し、公開を行う。また SHOE 及び関連ツールを用いた API 経由での AI モジュールとの連動を実現する。

C. 研究結果

I. 深層学習を用いた大規模遺伝子発現データベースからの重要遺伝子群の判別

近年、大規模かつ多次元的な遺伝子発現データが蓄積されつつあり、これらのデータを格納したデータベースが構築されている。特に、多様な要素が絡み合った、化合物毒性のメカニズムは極めて複雑である。Percellome 等の化合物毒性に関わる発現データベースは、この様な複雑性を内包している。したがって、Percellome から、有用な情報を抽出し活用するためには、この様なデータの複雑性を扱うことが出来る解析法が、必要不可欠である。しかしな

がら、この様な複雑性の高いデータから有用な情報を引き出すことは、従来の研究で用いられてきた統計解析法等では極めて困難である。

今まで、我々は、Deep Learning、および、その他の機械学習アルゴリズムを利用して、Percellome データベースから、一定の特徴量を認識して、有意な遺伝子を抽出するという問題を設置し、その実現可能性について検討を行ってきた。

Percellome データベースでは、化学物質投与下の遺伝子の発現情報を画像、または、数値で表す。これまで、Percellome データベースの解析では、化学物質投与下における有意な遺伝子を同定するために、長年経験とトレーニングを積んできた研究者が、それぞれの遺伝子の発現パターンの画像を目で見て、分類を行ってきた (図1)。

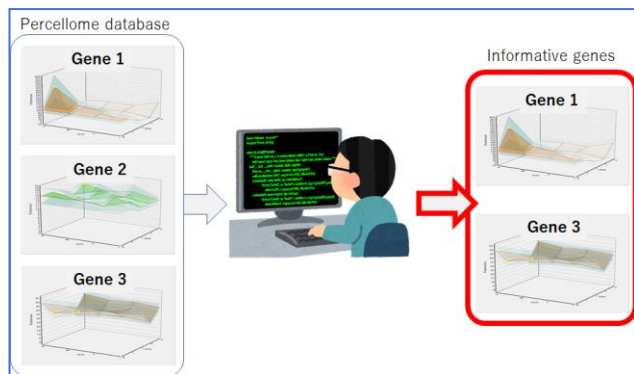


図1：専門家の visual inspection による、有意な遺伝子の同定

しかしながら、Percellome データベースは、100以上の化学物質それぞれに対し、40,000以上のプロンプの発現データが含まれており、400万以上の発現データを記述した画像で構成されている。この様な、大量の画像の全てを目で見て、有意な遺伝子を分類するためには、多大な労力と時間が必要となる。

ところで、近年、機械学習の手法の発展は目覚ましく、機械学習を用いた画像解析の手法も数多く開発されつつある。特に、deep learningの手法の発展はめざましく、分野によっては、人間の能力を超えつつある(David Silver et al. 2006, Nature 529)。医療分野で最も成功している例としては、deep learningの技術を使用した医療画像データ解析であり、非常に高精度に皮膚がんの分類を行うことのできる deep learning モデルが、発表されている(Andre Esteva et al. 2017, Nature 52)。また、最近では、発見が難しい、早期胃がんの自動検出を行う、deep learning モデル

が発表されている(Yoshimasa Sakai et al. 2018, 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society)。

そこで、昨年度 (H29年度)、我々は、Percellome データベースの画像データから、効率良く有意な遺伝子を判別する手法の構築を目的として、最新の多数の層で構成される deep learning algorithm を用いた、画像解析システムの実装を行った。実装した deep learning に基づく分類器の予測精度の検証を、熟練した研究者が目で見えて作成した画像データセットを用いて行った。また、Percellome データベースの数値データを用いた、deep learning 以外の機械学習による分類器も作成し、その予測精度を検証した。

その結果、画像を用いて deep learning で分類を行った場合の方が、数値を用いて deep learning 以外の機械学習を用いて分類を行った場合よりも、判別結果に偏りが少なく、より適切に有意な遺伝子群を分類できることが分かった (図2)。

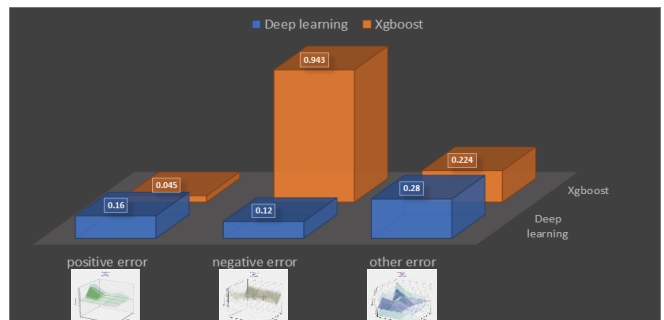


図2: deep learning に基づく分類器と、それ以外の機械学習モデル (Xgboost) に基づく分類器の、予測精度の比較：各分類器は、遺伝子を3つのクラス、positive class (化学物質により発現が up regulate されるクラス)、negative class (化学物質により発現が抑制されるクラス)、rough class (化学物質により up regulate も抑制もされないクラス)、に分類する。図のバーは、各分類器の、各クラスにおける、誤分類の割合を示している。青いバーに示すように、deep learning に基づく分類器では、各クラスの誤判定率は同様であった。一方、橙色のバーに示すように、deep learning 以外の機械学習に基づく (Xgboost) 分類器では、negative class の誤判定率が、94%と、非常に高い値を示した。

しかしながら、この様な deep learning に基づく分類器であっても、その予測精度は85%程度の正答率であり、実用化のためには、改良を加えてより高い精度の分類器を構築する必要がある。現在の我々の、deep learning に基づく分類器の精度が、85%の正答率にとどまっている原因としては、以下の3つの問題点が挙げられる。

問題点1：現在使用している画像データには、軸の情報や遺伝子名などの、ノイズの原因となる情報が含まれている（図3）。

問題点2：現在使用している画像データは、デフォルトの角度から見た、画像である。学習により適した角度がある可能性がある（図4）。

問題点3：現在使用している画像データはサイズが小さく（32×32ピクセル）、白黒である。

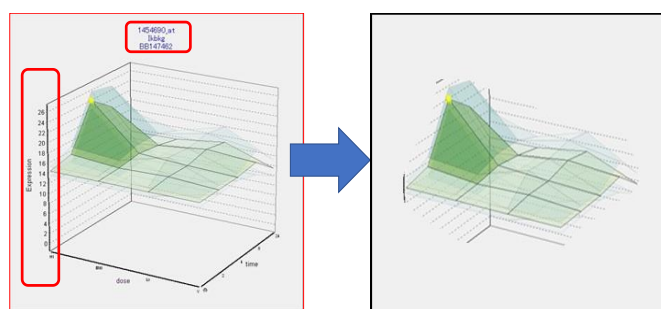
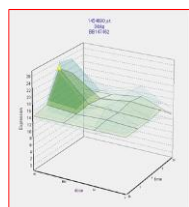


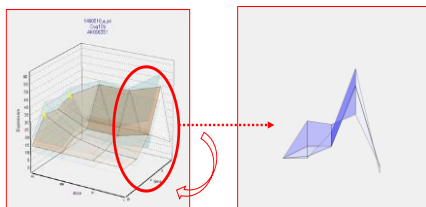
図3. ノイズの原因となる情報の削除：軸、遺伝子名、タイトルなどの余分な情報があり、これらが、予測精度を下げる原因となっていると考えられる。

H29年度の、我々のdeep learningモデルで使用した画像



デフォルトの角度からの画像

画像の回転



専門家は、遺伝子を分類するときには、画像を回転させて、見ている。

図4. 軸を回転させた画像の生成

これらの問題を解決するために、今年度（H30年度）、我々は、新しいトレーニングデータとなる画像セットを作成した。この画像セットには、軸の情報や遺伝子名などのノイズの原因となる情報が含まれない（図3）。更に、複数の角度から見た画像を生成し、複数のトレーニングデータを作成した。また、これらの画像セットの画像サイズを、500×500pixelとし、色をRGBカラーとした。これらのデータセットを用いて、deep learning モデルを精緻にチューニングし、分類器を作成した。その後、チューニングした分類器の精度の検証を行った。

画像データセットとしては今回、以下の4つの画像データセットを生成し、それぞれのデータセットを用いて、4つの deep learning に基づく分類器を構築した（図5, 6, 7, 8）。

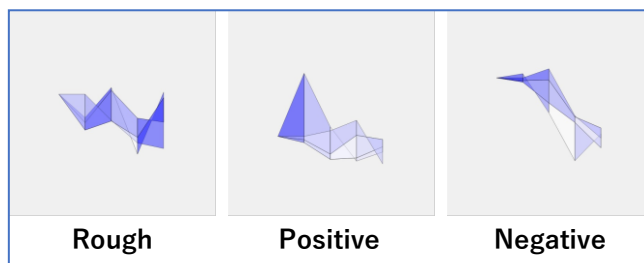


図5. 画像データ①:横の方向から見た画像で、遺伝子発現量の変動に関する P 値の情報を含まない画像: 500×500 ピクセルのサイズの、RGB カラー画像。各タイムポイント、化合物投与量下における、遺伝子発現量が有意に変化しているかどうかに関する情報 ($p < 0.05$) を含まない。Positive、negative、other クラスのデータとして、4,591、377、496 プロープに対応した画像データを作成した。

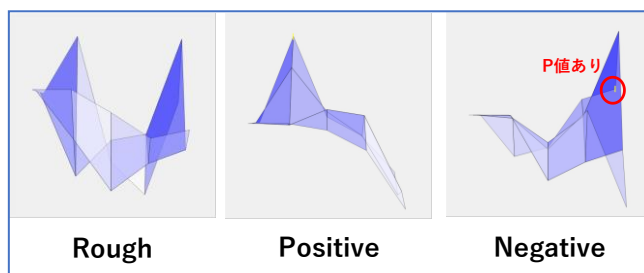


図6. 画像データ②:横の方向から見た画像で、遺伝子発現量の変動に関する P 値の情報を含む画像: 500×500 ピクセルのサイズの、RGB カラー画像。各タイムポイント、化合物投与量下における、遺伝子発現量が有意に変化しているかどうかに関する情報 ($p < 0.05$) を含む。Vehicle と比較して、あるタイムポイントと化合物投与量下において、遺伝子発現量が有意に変動をしている ($p < 0.05$) 場合、画像中にアスタリスクを表示した。Positive、negative、other クラスのデータとして、4,591、377、496 プロープに対応した画像データを作成した。

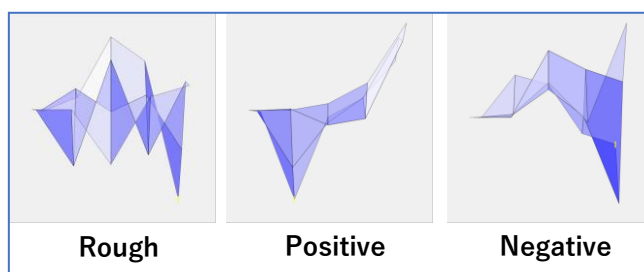


図7. 画像データ③:画像データ②を、上下方向に反転させた画像: 500×500 ピクセルのサイズの、RGB カラー画像。各タイムポイント、化合物投与量下における、遺伝子発現量が有意に変化しているかどうかに関する情報 ($p < 0.05$) を含む。Vehicle と比較して、あるタイムポイントと化合物投与量

下において、遺伝子発現量が有意に変動をしている ($p < 0.05$) 場合、画像中にアスタリスクを表示した。Positive、negative、other クラスのデータとして、4,591、377、496 プロープに対応した画像データを作成した。

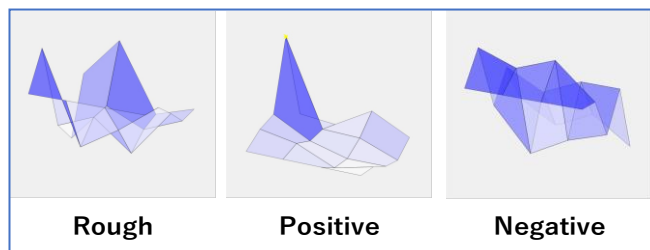


図 8. 画像データ④:デフォルトの角度から見た画像で、遺伝子発現量の変動に関する P 値の情報を含む画像: 500×500 ピクセルのサイズの、RGB カラー画像。各タイムポイント、化合物投与量下における、遺伝子発現量が有意に変化しているかどうかに関する情報 ($p < 0.05$) を含む。Vehicle と比較して、あるタイムポイントと化合物投与量下において、遺伝子発現量が有意に変動をしている ($p < 0.05$) 場合、画像中にアスタリスクを表示した。Positive、negative、other クラスのデータとして、4,591、377、496 プロープに対応した画像データを作成した。

deep learning に基づく、画像データから有意な遺伝子を予測する分類器の開発に際しては今回、deep learning モデルとして 18 層 (畳み込み層やプーリング層) のレイヤーで構成される、resnet モデルを用いた(Kaming He et al. 2015, arXiv:1512)。Resnet は、画像分類で高い精度を示すことが多く、ImageNet2015 (<http://image-net.org/challenges/LSVRC/2015/>) における画像分類に関するコンペティションで、最も高い精度を示したモデルである(Kaming He et al. 2015, arXiv:1512)。今回、我々が作成した分類器は、Resnet の多数の層を活用して、画像データから、特徴量の自動抽出を行う。その後、最後の層において、抽出した特徴量に基づき、それぞれの遺伝子を、3つのクラスに分類する。3つのクラスは、positive class (化学物質により発現が up regulate されるクラス)、negative class (化学物質により発現が抑制されるクラス)、rough class (化学物質により up regulate も抑制もされないクラス)、である (図 5, 6, 7, 8)。Positive と negative class に分類された遺伝子が、有意な遺伝子と予測されるものである。

今回、上記の画像データ①から④を用いて、4つの deep learning モデルに基づく、4つの分類器を構築した。構築した分類器に対して、5 fold cross validation を実施し、予測精度を検証した。図 9 に、各分類器の予測精度の検証結果を示す。

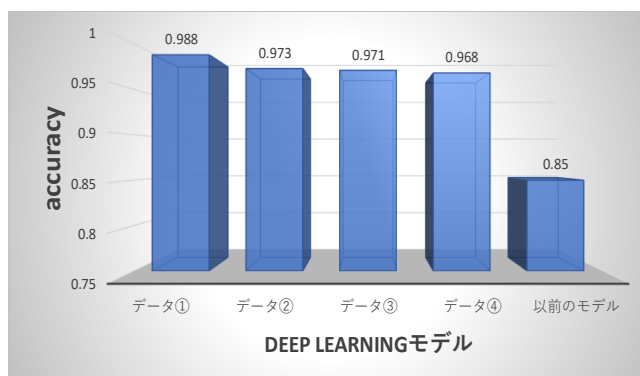


図 9. 予測精度: 各分類器の予測精度の比較を行った。それぞれの予測モデルの予測精度、正答率、をバーで表した。画像データ①を基にトレーニングを行った分類器の正答率は、98.8%と、非常に高い値を示した。

図 9 のように、データ①から④を基にトレーニングした分類器の予測精度は、先行研究の分類器の予測精度よりも、10%以上正答率が高く、95%以上の正答率という非常に良い予測精度を示した。以上の結果から、画像データからノイズとなる情報を削除し、3D 描画を適切な角度に調整することで、分類器の予測精度が大幅に改善されることが示された。

II. Garuda Platform 上での Percellome と SHOE の連動

Garuda Platform 上で、Percellome DB と SHOE の連動をより密にした。図 10 に連動の概要を、そのあとに、一連の画面を示す。

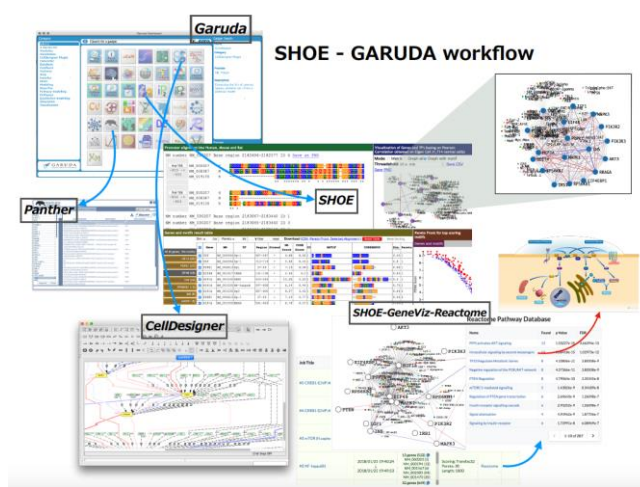


図 10. ソフトウェアの連動概要

2. Integration of ChIP-seq Array data into SHOE

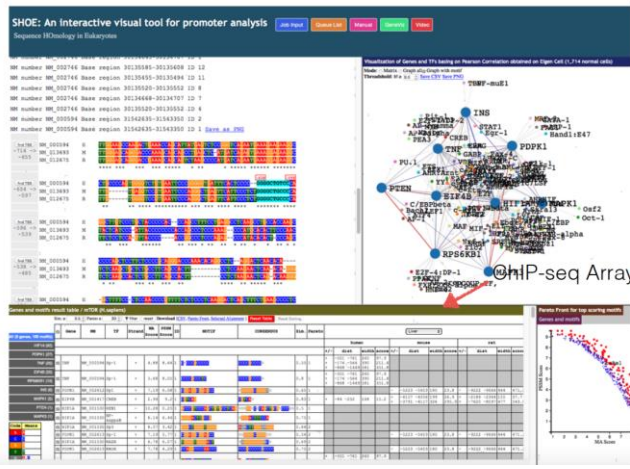


図 1 1. ChIP-seq アレイデータを導入した SHOE 画面

3. Integration of ChIP-seq Array data into SHOE

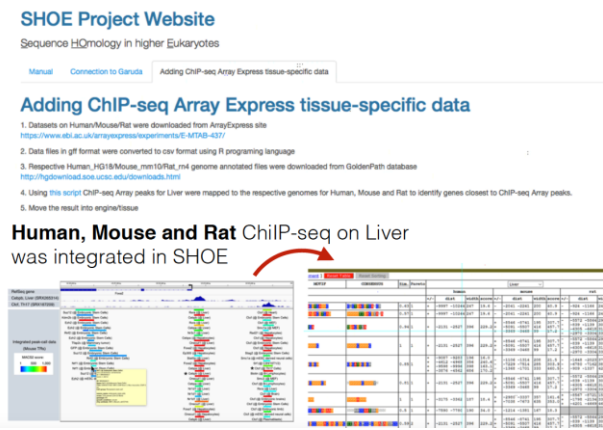


図 1 2. SHOE への ChIP-seq アレイデータの統合

Example 1 of ChIP-seq peak in Liver predicted by SHOE

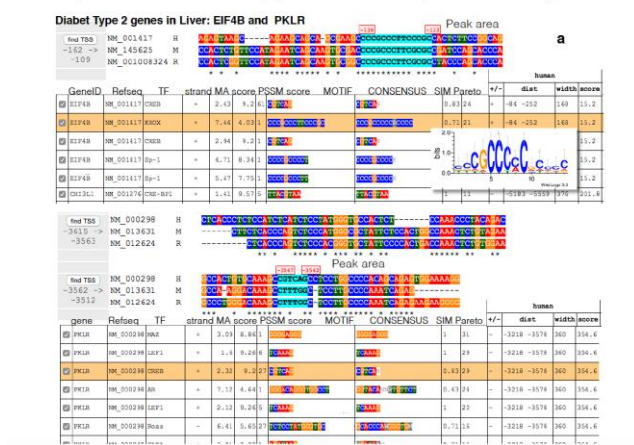


図 1 3. 解析結果例

D. 考察

深層学習は、Percellome データベースのような大規模な発現データの自動分類に有効であることがわかった。今後は、未知のデータセットに対し分類器を適用し予測を行い、予測結果を、専門家の目視による結果と比較する。これにより深層学習による分類結果と専門家の目視による分類結果との差異を基に、さらに、より精緻な分類器のチューニングが可能と考えられる。

E. 結論

今回の成果で、人手ではできなかった大規模の遺伝子発現と毒性への連動解析が可能となる。また、SHOE と Percellome の連動も密になり、よりスムーズな解析が可能となった。

今後は、この成果をフィードバックし、さらに高速かつ高精度の解析プロセスの展開を目指す。

F. 研究発表

1. 論文発表

- ① N Polouliakh. In Silico Transcription Factor Discovery via Bioinformatics Approach: Application on iPSC Reprogramming Resistant Genes. Leveraging Biomedical and Healthcare Data, 2019 – Elsevier
- ② Polouliakh N, Horton P, Shibani K, Takata K, Ludwig V, Ghosh S, Kitano H. Sequence homology in eukaryotes (SHOE): interactive visual tool for promoter analysis. BMC Genomics. 2018 Sep 27;19(1):715. doi: 10.1186/s12864-018-5101-3.

2. 学会発表

- ① Natalia Polouliakh. Garuda/Shoe and Percellome analytic workflow. OPENTOX ASIA 2018, Asahi Seimei Otemachi Building, Tokyo, May 24, 2018.
- ② Takeshi Hase, Samik Ghosh, Ken-ichi Aisaki, Satoshi Kitajima, Jun Kanno, and Hiroaki Kitano. DTOX: Deep neural network-based computational framework to analyze omics data in Toxicology. OPENTOX ASIA 2018, Asahi Seimei Otemachi Building, Tokyo, May 25, 2018.
- ③ Kitano, H. AI-driven systems toxicology. Toxicogenomics for Accelerated and Refined Hazard Identification of Chemicals (Joint Symposium between IUTOX and IUPHAR), 18th World Congress of Basic and Clinical Pharmacology, Kyoto International Conference Center, Kyoto, July 6, 2018.

(invited)

- ④ 北野宏明. AI 駆動型システム医科学の展望. 山口大学第2回シンポジウム: 人工知能・システム医学による難治性疾患への新たな挑戦, KKR 山口あさくら, 山口, Aug. 4, 2018. (invited)
- ⑤ Kitano, H. AI Grand Challenges. 2018 年日台科学技術フォーラム, リージェント台北, 台湾, Sep. 10, 2018. (invited)
- ⑥ 北野宏明. 人工知能駆動型システム毒性学の展望. CBI学会 2018 年大会, タワーホール船堀, 東京, Oct. 9, 2018. (invited)
- ⑦ 北野宏明. AI 戦略とムーンショット. データサイエンティスト協会シンポジウム, JP タワーホール&カンファレンス東京, 東京, Oct. 19, 2018. (invited, keynote)

G. 知的財産権の出願・登録状況

なし