

## 二国間交流事業 共同研究報告書

平成23年4月8日

独立行政法人日本学術振興会理事長 殿

共同研究代表者所属・部局 京都大学・化学研究所

職・氏名 (ふりがな) 准教授・五斗 進 ごとう すずむ

1. 事業名 相手国（フランス）との共同研究 振興会対応機関（ INSERM ）

2. 研究課題名 癌のバイオインフォマティクスにおけるアルゴリズムとデータベースの開発

3. 全採用期間

平成21年4月1日～平成23年3月31日（2年 ヶ月）

4. 研究経費総額

(1) 本事業により交付された研究経費総額 5,000 千円

初年度経費 2,500 千円、 2年度経費 2,500 千円、 3年度経費   千円

(2) 本事業による経費以外の国内研究経費総額 1,000 千円

5. 研究組織

(1) 日本側参加者

氏名 (ふりがな)	所属・職名	研究協力テーマ
ごとう すすむ 五斗 進	京都大学化学研究所・准教授	データベース構築および研究の総括
かねひさみのる 金久 實	京都大学化学研究所・教授	データベース構築と改良
たむら たけゆき 田村 武幸	京都大学化学研究所・助教	解析手法の開発と改良
ほんだ わたる 本多 渉	京都大学理学研究科・大学院生	解析手法の改良
みずたに さやか 水谷 沙弥佳	京都大学理学研究科・大学院生	解析手法の改良とデータ解析
じょん ぶらうん John Brown	京都大学情報学研究科・大学院生	解析手法の改良
たなか みちひろ 田中 道廣	京都大学薬学研究科・ポスドク研究員	データベースの改良

(2) 相手国側研究代表者

所属・職名・氏名 キュリー研究所・上席研究員・Jean-Philippe Vert

(3) 相手国参加者（代表者の氏名の前に○印を付すこと）

氏名	所属・職名（国名）	研究協力テーマ
○Jean-Philippe Vert	キュリー研究所・上席研究員（仏）	研究の総括
Yoshihiro Yamanishi	キュリー研究所・研究員（仏）	解析手法の開発と改良
Kevin Bleakley	キュリー研究所・ポスドク研究員（仏）	解析手法の改良
Fantine Mordelet	キュリー研究所・大学院生（仏）	解析手法の改良とデータ解析
Mikhail Zaslavskiy	キュリー研究所・大学院生（仏）	解析手法の改良
Brice Hoffmann	キュリー研究所・大学院生（仏）	解析手法の改良

## 6. 研究概要（研究の目的・内容・成果等の概要を簡潔に記載してください。）

本研究の目的は、様々なハイスループットデータから癌関連遺伝子を同定するだけでなく、癌関連遺伝子が関わる分子間相互作用を表す癌関連パスウェイや、癌特異的なパスウェイモジュールを網羅的に予測する新しいアルゴリズムやデータベースを開発することである。また、癌関連遺伝子の環境因子との相互作用や医薬品応答性を解析し、癌に対する分子標的薬開発に向けたケモゲノミクス手法の開発や癌の分子メカニズムに対して生物学的な考察を行うことも、本研究の大きな目的である。

京大チームでは、主にデータベースの開発と解析を担当した。データベース開発に関しては、癌関連遺伝子の情報や癌関連遺伝子が関わる分子間相互作用ネットワークを表すパスウェイ情報のデータベースを構築した。これらは京大チームが開発してきた KEGG PATHWAY データベースにおける Human diseases: Cancers セクションの改良という形で実現し、公開している。また、癌関連遺伝子とともに環境因子、診断マーカー、治療薬の情報をまとめた KEGG DISEASE データベース、医薬品の化学構造、ターゲット、代謝酵素、効能情報などをまとめた KEGG DRUG データベースに関しても癌に関する情報を随時追加し、公開している。

キュリー研チームでは、主にアルゴリズムの開発と解析を担当した。京大チームが構築したデータベースを用いて新たな癌関連パスウェイを同定するために、アレイ CGH における DNA コピー数情報とマイクロアレイにおける遺伝子発現情報を同時に解析するアルゴリズムを開発した。具体的には、スパース正準相関分析という統計手法を応用することによって二つの異質な高次元プロファイルの同時解析を実現した。また、遺伝子発現情報を癌組織と通常組織間で比較し、癌特異的なパスウェイモジュールを検出するアルゴリズムも開発した。サポートベクタマシンと呼ばれる分類のための統計モデルの中で、京大チームが構築したデータベースの情報を用いている点が特徴である。

医薬品データの解析に関しては、抗癌剤を含む医薬品の化学構造、薬効、副作用の相関解析を行った。京大チームが開発している KEGG DRUG とゲノムネット医薬品データベースから、化学構造情報と薬効・副作用情報とをそれぞれ入手して利用した。薬効と副作用情報に関しては自然言語で書かれた添付文書からテキストマイニングの手法を用いることにより抽出した。キュリー研グループが中心となって、化学構造から薬効や副作用を予測する回帰モデルを開発した。両チームが連携してこのモデルをネットワーク推定法に応用することによって、薬効類似性から医薬品の標的タンパク質を予測するアルゴリズムおよび副作用の類似性から標的タンパク質を予測するアルゴリズムの開発に成功した。

医薬品の代謝を含むネットワークの解析に関しては、影響度の理論的な研究を進めた。影響度とは、代謝ネットワークにおいて、ある反応を阻害した際に、その反応も含めて全部でいくつの反応が影響を受けるかを示す指標である。本研究では、本指標をネットワークが閉路を含む場合にも対応できるように拡張し、影響度を計算するための効率的なアルゴリズムを開発した。また、ランダムグラフなどを用いてモデル上での影響度の平均や分散などを理論的に計算した。さらに、KEGG PATHWAY のデータを用いた実験結果と比較して、理論の妥当性を評価した。