

令和 4 年 6 月 2 日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K12215

研究課題名(和文) 腫瘍組織多様性理解のための単一細胞に対する時空間内包型数理解析手法の提案

研究課題名(英文) Proposal of a Spatio-Temporal Inclusionary Mathematical Analysis Method for Single Cells to Understand Tumor Tissue Diversity

研究代表者

片山 琴絵 (Katayama, Kotoe)

東京大学・医科学研究所・准教授

研究者番号：40581195

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本研究は単一細胞解析において現在のほとんどの技術が細胞の破壊を必須としており、その時点での細胞情報を得ることはできても細胞自体の時間および空間情報の欠失がもたらす問題を数理的モデルを用いて解決しようとするものである。これまでの手法でプロセスの異なる複数個の細胞を区別なく同一集団として扱うため、がんの発生メカニズムや転移といった現象を特徴付けるのに限界がありこれを数理的な観点から記述を定義を行い既存手法との比較見当を行なった。また単一細胞における細胞周期による揺らぎを確率分布として記述することにより腫瘍組織内のがん細胞の多様性を明らかとするモデルの構築を行った。

研究成果の学術的意義や社会的意義

単一細胞解析において細胞の破壊を必須としており、時間および空間情報が欠失する問題を数理的モデルを用いて解決しようとするものである。腫瘍組織内のがん細胞の多様性が「腫瘍内不均一性」としてがんの発生には、突然変異した細胞のクローンに起因するプロセスと、がん幹細胞から分化したプロセスの両方が関与していると考えられている。これまでの手法では、複数個の細胞を区別して扱うことができなかったがこれを数理的な観点から細胞周期による揺らぎを確率分布とすることで腫がん細胞の多様性を明らかとするモデルの構築を行った

研究成果の概要(英文)：This study attempts to solve the problem that most of the current techniques for single cell analysis require the destruction of cells, and although it is possible to obtain cellular information at that point in time, the temporal and spatial information of the cells themselves is lost. Since conventional methods treat multiple cells with different processes as a single population without distinction, they are limited in their ability to characterize phenomena such as cancer development mechanisms and metastasis. In addition, we constructed a model that clarifies the diversity of cancer cells in tumor tissue by describing fluctuations in the cell cycle of a single cell as a probability distribution.

研究分野：バイオインフォマティクス

キーワード：がん バイオインフォマティクス 単一細胞解析 シンボリックデータ 接合分布関数

1. 研究開始当初の背景

遺伝子が発異することによって引き起こされる病気であるがんは、NGS から出力された全ゲノムシーケンズデータやトランスクリプトームデータに基づき、様々な解析が実施され、原因となる遺伝子の発見や創薬につながるがん抑制遺伝子などが発見されつつある。これらの研究成果を元にして近年開発が進んでいる「分子標的薬剤」は、一時的には有効であるが、薬剤耐性を有するがん細胞の出現により多くの患者は再発することから、がんの治療成績への寄与は、不十分と言わざるを得ない。一方、近年開発された免疫チェックポイント阻害薬が、多くのがん種において、顕著な臨床効果を示し、末期のがん患者においてもしばしば治癒をもたらすことが報告されるようになってきた。このように大まかな生体内共通の特徴を観測してはいるが、効果や予後に極めて大きな差があり、その作用の全容が未だ解明できていない。個別化医療を見据え、この作用の解明は腫瘍を平均的に観測することでは解明できないと考え、単一細胞解析データによる腫瘍構成細胞の多様性と関連性を無視したものであるのかを確認する必要があると考えた。NGS データ解析により、多くのがん種で、免疫に関わる遺伝子の異常が生じており、これらの異常には、染色体の欠失や逆位、転座、重複などさまざまなタイプの遺伝子構造異常が含まれていたが、すべての例で、タンパク質に翻訳されない 3'非翻訳領域と呼ばれる領域の欠損と PD-L1 遺伝子の発現の顕著な上昇が認められたことが明らかとなっている [Ogawa S. et al. (2016) Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers. *Nature*.] このように、染色体構造異常と遺伝子発現の上昇という共通の特徴が見られるにも関わらず、治療効果には大きな差が見られている。遺伝子異常数の多いがんの方が高い抗腫瘍効果が得られるという報告がなされている一方で、遺伝子異常数の過多で治療効果の 2 群比較を行ったところ、差がなかったという報告もある [Carbone D, et al. (2017) First-Line Nivolumab in Stage IV or Recurrent Non-Small-Cell Lung Cancer. *NEJM*.]。これらは複数の細胞腫が混在する腫瘍細胞集団を標的とした治療が確立されても、別の集団に対して腫瘍内不均一性を理由としてそれらが有効でない著例である。例えば乳がんの単一細胞解析において多数のサブクローナルおよび de novo 変異が見出され、点突然変異が徐々に進化したことを示唆する結果や [Wang Y, et al. (2014) *Nature*.]、多発性骨髄腫における単一細胞解析では 2 つの異なるクローンが独立して同じ収束表現型を獲得する並列進化パターンを介して病気が発症することが実証された [Melchor L, et al. (2014) *Leukemia*.]。このような現状から、現在共通だと思われる特徴は NGS から得られた細胞平均的な特徴の一面を表したものであり、生体内個別細胞特異的な解析により薬理作用の全容を明らかにできると考えたが、これを実現可能にするために細胞周期の揺らぎを取り入れた複合データ情報処理を可能にするデータ解析手法が存在せず、本研究でそれらを構築しようと考えた。

2. 研究の目的

生命科学の研究方法として、「単一細胞 (シングルセル) 解析」が大きな注目を浴びている。これは細胞を個別に解析していく研究方法であり、細胞集団を解析して平均値を捉えるこれまでの手法では見逃されてきた新しい細胞種、また、これまでになくレベルでの細胞機能を解明できるようになった。単一細胞解析の応用が期待される重要な分野の 1 つががんゲノミクスである。腫瘍組織内のがん細胞の多様性が「腫瘍内不均一性」として注目されており、細胞間での著しい異質性の理解が求められている。がんの発生には、突然変異した細胞のクローンに起因するプロセスと、がん幹細胞から分化したプロセスの両方が関与している [Clevers H. (2011) *Nat Med*] と考えられているが、これまでの手法では、プロセスの異なる複数個の細胞を区別なく同一集団として扱うため、がんの発生メカニズムや転移といった現象を特徴付けるのに限界がある。また、複数の細胞腫が混在する腫瘍細胞集団を標的とした治療が確立されても別の集団に対して腫瘍内不均一性を理由として有効でない可能性があるため、治療の失敗および再発の要因として細胞個別の特徴抽出はますます重要視されている。単一細胞解析で行われる一般的な方法は、腫瘍細胞集団を単一細胞に分解したのちに RNA-seq を行い、各細胞で得られた遺伝子発現量をクラスタリングすることで同様の特性を有する細胞をグループ化し、腫瘍細胞集団像を再構築する。得られた腫瘍細胞像から、細胞集団全体で共有している高発現遺伝子群や逆に、これまで知られていなかった未知の亜集団で特異的に発現している遺伝子を解析することで、がんの理解、腫瘍マーカーの同定、創薬に繋げようとするものである。腫瘍細胞集団像を再構築する手法として、古典的なクラスタリング方法に加えて、単一細胞に特化した多くの手法が提案されている。ACCENSE [Shekhar K, et al. (2014) *Proc Natl Acad Sci USA*.] は、クラスタ数を指定する必要なしに次元圧縮アルゴリズム t-SNE と密度ベース分類を組み合わせる。また SNN-Cliq [Xu C, Su Z (2015) *Bioinformatics*.] ではグラフ理論に基づくアルゴリズムによってクラスタリングを達成する。クラスタリングは基盤となる腫瘍細胞のグループ構造を特定の時点で観測された姿として明らかにするが、各グループが細胞集団の初期化から存在したのか、それともある時点で分化したのか、といった系統関係に関する情報を得ることはできない。これに対する 1 つの挑戦として分岐理論を用いて遺伝子発現多様性をモデル

化する SCUBA [Marco E, et al. (2014) Proc Natl Acad Sci USA.]があるが、適用には時間的情報が必要であり、これは実験的に得ることが困難なことが多い。ここで認識しておかなければならないことは、単一細胞解析の発現量を元にクラスタリングされた腫瘍細胞像において、同一グループに属するとされた細胞はすべて同じ時点で出現したわけではないということである。発現量から細胞集団内で類似性が高いと判断され同じグループに属すると判定された細胞であっても、その発現量は細胞周期によって一定の揺らぎをもって変動している [Sasagawa Y, et al. (2013) Genome Biology]。また単一細胞解析では発現量だけではなく、細胞周期全ゲノム/エキソソーム単一細胞配列決定法 (nuc-seq) などによって細胞個別の遺伝子変異を検出する実験的手法が確立されつつあるが、これも発現量と同様に細胞周期によって一定の揺らぎが存在する。しかしながらこれまで、細胞周期の揺らぎを考慮して腫瘍組織内のがん細胞の多様性とその相互作用を解明しようとする取り組みはなされてこなかった。このような学術的背景のなかで本研究課題の目的は、単一細胞内での規則的な細胞周期によって引き起こされる確率的ノイズなどの変動を前提とした遺伝子発現量と遺伝子変異を統合分析するためのデータ解析手法の提案と実行によって、腫瘍組織内のがん細胞の多様性の起源と進化を明らかとし、患者特異的治療や創薬への応用を可能とすることである。

3. 研究の方法

これまでの単一細胞分析ではある一時点で得られたデータから時系列推移を類推し、細胞の進化を明らかにしようとするアプローチがほとんどであった。これは単一細胞分析のためのほとんどの技術は、細胞の破壊を必要とするため、時間情報は失われることに起因する。同様に、組織から単細胞を単離することは、空間的状况に関する情報の損失をもたらす。これら時間情報、空間情報の欠失は、細胞相互作用に大きく依存することが知られている腫瘍の進行を明らかにしようとする場合、特に問題である。本研究では単一細胞における細胞周期による揺らぎを確率分布として記述することにより腫瘍組織内のがん細胞の多様性を明らかとした。近年の情報環境の発展により、収集されるデータが複雑化しつつあり、そこから有益な情報を適切に抽出するための解析手法の開発が期待されている。しかし既存の方法では複雑なデータを適切に扱うことが難しい事例も多い。そのような背景を契機として、シンボリックデータ解析が Diday (1988) により提唱された。シンボリックデータ解析法における重要な特徴は、解析対象の記述の柔軟性である。一般的な多次元データ解析において、解析の対象となるデータは複数の個体であり、各個体は単一の値、もしくはベクトルの形で記述されることが多い。これに対して、シンボリックデータ解析法では個体の集まりそのものを実際の解析対象とし、その記述として、区間、確率分布など、従来の手法では直接扱えない形を考えることができる。申請者はすでに確率分布を対象とした次元縮小とクラスタリングに関するシンボリックデータ解析手法の構築を行なっている [Katayama K, et al (2010) Journal of the Japanese Society of Computational Statistics など]。本研究では単一細胞における細胞周期による揺らぎを確率分布として記述し解析対象とすることで情報エントロピーの取得が可能となり、腫瘍組織内のグループ構造をモデル化することができる。さらに、この手法ではこのグループ構造自身もまた混合分布によって記述される。これによりデータの種別 (遺伝子発現量と遺伝子変異) に構成されたグループ構造に基づいて [目的 2]: 単一細胞から得られた遺伝子発現量と遺伝子変異の統合的理解と、細胞間ネットワークの構築を行った。これは確率分布として記述された遺伝子発現量と遺伝子変異をコピュラ (接合分布関数) で評価することで実現する。コピュラは同時確率分布と周辺分布を結合するものである。確率分布で記述された遺伝子発現量と遺伝子変異との関係が線形的に表すことができず、相関係数がゼロであったとしても両者の間に相互関連は存在するのであり、その関係性を表す尺度としてコピュラを用いることでネットワークの構築を行った。コピュラでの評価は確率分布で記載された解析対象すべてに及ぶため、すべての解析対象は確率分布で記述されるとしている本研究課題の問題設定において、同一細胞内の発現量と遺伝子変異の関係性、腫瘍細胞組織内のグループ間関係性、また腫瘍組織細胞と正常組織細胞の関係性といった生体内システム全体の関係性を確率分布的ネットワークとして記載することが可能となった。

4. 研究成果

モデル化された腫瘍組織内確率空間モデルによって腫瘍細胞全体の細胞周期から逸脱して進化する単一がん細胞の進化および消滅過程を同定する。この目的のため細胞集合の平均的特性として、既存の Pathway、OMIM、Pubmed などから得た特徴量を情報エントロピーとすることで細胞集合間の距離を定義し、定量化する手法を提案した。得られた情報エントロピーは生体情報分布の距離として定義し、細胞間ネットワークと腫瘍組織内での情報関連性および集約を表現するものである。情報エントロピーが最大となるとコピュラで表現される確率分布間の関連性も最大となるような部分ネットワーク同士の連結を行うことで、多重クラスターからなる腫瘍細胞全体のネットワークを推定することができ、腫瘍内不均一性に関わるがん細胞の多様性が及ぼす影響をシミュレーションによって明らかとなった。以上のような数理的解析手法は生体システムの解明および医療・創薬に有用な情報抽出のためのツールとして重要である。

5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 7件/うち国際共著 1件/うちオープンアクセス 2件）

1. 著者名 Omori, Satotaka, et al.	4. 巻 32
2. 論文標題 Generation of a p16 Reporter Mouse and Its Use to Characterize and Target p16high Cells In Vivo	5. 発行年 2020年
3. 雑誌名 Cell Metabolism	6. 最初と最後の頁 814 ~ 828.e6
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.cmet.2020.09.006	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Fujimoto, Kosuke, et al.	4. 巻 -
2. 論文標題 Functional Restoration of Bacteriomes and Viromes by Fecal Microbiota Transplantation	5. 発行年 2021年
3. 雑誌名 Gastroenterology	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1053/j.gastro.2021.02.013	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Murakami Michio, Miura Fuminari, Kitajima Masaaki, Fujii Kenkichi, Yasutaka Tetsuo, Iwasaki Yuichi, Ono Kyoko, Shimazu Yuzo, Sorano Sumire, Okuda Tomoaki, Ozaki Akihiko, Katayama Kotoe, Nishikawa Yoshitaka, Kobashi Yurie, Sawano Toyoaki, Abe Toshiki, Saito Masaya M., Tsubokura Masaharu, Naito Wataru, Imoto Seiya	4. 巻 -
2. 論文標題 COVID-19 risk assessment at the opening ceremony of the Tokyo 2020 Olympic Games	5. 発行年 2021年
3. 雑誌名 Microbial Risk Analysis	6. 最初と最後の頁 100162 ~ 100162
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.mran.2021.100162	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Tsuda Yusuke, Hirata Makoto, Katayama Kotoe, ... & Miyano Satoru, Imoto Seiya, Shibata Tatsuhiro, Nakagawa Hidewaki, Yamaguchi Rui, Tanaka Sakae, Matsuda Koichi	4. 巻 145
2. 論文標題 Massively parallel sequencing of tenosynovial giant cell tumors reveals novel CSF1 fusion transcripts and novel somatic CBL mutations	5. 発行年 2019年
3. 雑誌名 International Journal of Cancer	6. 最初と最後の頁 3276 ~ 3284
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/ijc.32421	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hirata Makoto, Asano Naofumi, Katayama Kotoe, Yoshida Akihiko, Tsuda Yusuke, Sekimizu Masaya, ... & Imoto Seiya, Miyano Satoru, Kawai Akira, Yamaguchi Rui, Ichikawa Hitoshi, Matsuda Koichi	4. 巻 10
2. 論文標題 Integrated exome and RNA sequencing of dedifferentiated liposarcoma	5. 発行年 2019年
3. 雑誌名 Nature Communications	6. 最初と最後の頁 1-12
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41467-019-13286-z	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Matsuno Yusuke, Atsumi Yuko, Shimizu Atsuhiko, Katayama Kotoe, Fujimori Haruka, Hyodo Mai, Minakawa Yusuke, Nakatsu Yoshimichi, Kaneko Syuzo, Hamamoto Ryuji, Shimamura Teppai, Miyano Satoru, Tsuzuki Teruhisa, Hanaoka Fumio, Yoshioka Ken-ichi	4. 巻 10
2. 論文標題 Replication stress triggers microsatellite destabilization and hypermutation leading to clonal expansion in vitro	5. 発行年 2019年
3. 雑誌名 Nature Communications	6. 最初と最後の頁 1-13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41467-019-11760-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 VanderWeele David J., Finney Richard, Katayama Kotoe, Imoto Seiya, Yamaguchi Rui, Wheeler David, Lack Justin, ... & Vander Griend Donald, Kubo Michiaki, Ratain Mark J., Miyano Satoru, Nakagawa Hidewaki	4. 巻 5
2. 論文標題 Genomic Heterogeneity Within Individual Prostate Cancer Foci Impacts Predictive Biomarkers of Targeted Therapy	5. 発行年 2019年
3. 雑誌名 European Urology Focus	6. 最初と最後の頁 416 ~ 424
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.euf.2018.01.006	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------