

日本医療研究開発機構 創薬基盤推進研究事業 事後評価報告書

I 基本情報

研究開発課題名： ゲノム情報科学と機械学習に基づく高機能抗体のデザイン開発技術に関する研究
Study of designing technology of high performance antibodies based on clinical genome data science
and machine learning

研究開発実施期間：平成 30 年 5 月 1 日～令和 5 年 3 月 31 日

研究開発代表者 氏名：加藤 洋人
Katoh, Hiroto

研究開発代表者 所属機関・部署・役職：
国立大学法人東京大学・大学院医学系研究科衛生学分野・准教授
The University of Tokyo, Graduate School of Medicine, Associate Professor

II 研究開発の概要

本研究開発計画の目的は、腫瘍等のヒト疾患組織に存在する B 細胞クローンあるいは特異的抗原を認識するヒト B 細胞クローン等に着目したシングルセル解析を含む詳細な免疫レパトアシーケンスを実施し、それらの免疫レパトア配列情報を対象とした情報科学を基盤として、抗体医薬品として開発可能な高機能抗体クローンをアミノ酸配列レベルで推定可能な情報技術の基盤を開発することであった。抗体医薬品の開発においては、アミノ酸改変によって高機能性クローンを探索する必要性が生じ、時間とコストがともに大きく問題となっている。より効率の良い抗体医薬品開発の実現のためには、科学的理論付けを背景としたアミノ酸改変デザイン技術の開発が必要とされている。

本研究開発計画では、おもにヒト腫瘍免疫環境をモデルケースとして B 細胞集団のシングルセル・シーケンスを含めた網羅的免疫レパトア解析を実施することによって、腫瘍特異的免疫グロブリン・クローンの重鎖と軽鎖の組合せ情報を大量に取得し、生体内の affinity maturation 過程で獲得したアミノ酸変化を高精細に捉えるとともに、そのような成熟プロセスを推定する高精度な計算科学的アルゴリズムを見いだすことを計画していた。

本研究開発計画は以下 4 つの研究開発項目によって進められ、各々以下の成果が得られた。

- (1) シングルセル解析を含む包括的免疫レパトアシーケンスデータの取得
ヒト臨床組織に浸潤する B 細胞クローンおよび先行研究によって新規のがん抗原であることが明らかにさ

れていた高硫酸化グリコサミノグリカン(高硫酸化糖鎖)を認識する B 細胞クローンに対して、シングルセル解析を含む詳細な免疫グロブリン・レパトアシーケンスを実施した。腫瘍等のヒト疾患環境特異的な免疫グロブリンクローンあるいはがん抗原特異的な免疫グロブリン・クローンの配列情報を大量に取得し、計算科学的アプローチによってそれらの affinity maturation 過程のアミノ酸変化を高精細に捉える基盤を構築した。シングルセルレパトア解析では、免疫グロブリンの重鎖配列と軽鎖配列の組み合わせを 1 細胞レベルで確定することが可能であり、affinity maturation 過程における重鎖・軽鎖のアミノ酸変化プロセスを高精度に取得できる。このような多様な臨床組織から取得した高精細な免疫レパトア配列情報のデータベースは、汎用性と発展性のある創薬基盤技術の開発基盤となる重要なデータベースとなる。

(2) ゲノムデータサイエンスによる高機能抗体のデザイン設計技術の開発

上記(1)で同定されたヒト疾患環境特異的な免疫グロブリン・クローンおよび硫酸化糖鎖を認識するヒト B 細胞クローンのアミノ酸配列情報を用いた計算科学によって、それらに特徴的な抗体アミノ酸配列の特性およびその成熟プロセスにおける変化の過程を抽出した。研究期間を通して、ヒト免疫グロブリン分子に関するアミノ酸配列の特性を抽出するアルゴリズムの開発を進めることができた。成果の一部として、免疫レパトアデータに機械学習法を適応することに成功し、免疫グロブリン配列情報のみを用いて、その由来組織ががん組織か非がん部組織かを識別するアルゴリズムの開発に成功した。ヒト疾患組織に存在する B 細胞および抗原特異的な B 細胞から取得した豊富なヒト抗体配列情報を基盤とする計算科学によって高機能抗体分子のデザイン化の基盤とするという独自性が高い研究を進めることができたといえる。これらの成果は、効率良い抗体医薬品開発という社会的ニーズに応える新技術の創出に資する基盤的成果である。

(3) 抗体分子の発現・精製

上記(1)および(2)によって同定・改変されたヒト腫瘍環境特異的な免疫グロブリンあるいは高硫酸化糖鎖を認識するヒト免疫グロブリンの配列情報を基盤として、実際の低分子化抗体等としての発現・精製を実施し、それらの抗体分子の安定発現量の評価等、機能性の検討を進めた。それらの実験データは上記(2)の研究開発項目へと順次還元され、計算科学的アルゴリズムの洗練化に生かされた。その結果、一部の抗体分子については、上記(2)等に基づくアミノ酸改変によって可溶性発現量を増加させることが可能であることが明らかになった。すなわち、ヒト生体内における affinity maturation 過程のアミノ酸変化の計算科学的解析を元に、抗体分子の高機能化がデザイン可能だという Proof-of-concept の一部を示すことができたといえる。

(4) 高機能抗体の生物学的機能検証と機械学習へのフィードバック

上記(1)～(3)で同定・デザインされた抗体分子について、多様なヒト抗体配列についての機能的検証を進めた。腫瘍特異的なヒト抗体クローンおよび高硫酸化糖鎖がん抗原を認識するヒト抗体クローン、およびそれらのアミノ酸改変抗体クローンについて、種々のがん細胞に対する結合性やマウスモデルにおける抗腫瘍効果の検証などを進めた。これらのアミノ酸改変抗体に関する機能評価については上記(2)の研究開発項目へも還元されることで、計算科学アルゴリズムの洗練化が進められた。その結果、抗体医薬品として応用可能性のある複数の高機能抗体を同定・設計することに一部で成功している。また興味深いことに、ヒト腫瘍環境から単離された抗硫酸化糖鎖抗体が、低 pH 依存性に抗原との反応性が明瞭に増加することが新規に明らかになり、さらにそのような生化学的特性を生む抗体アミノ酸配列の特性の一端を見出した。腫瘍環境は低 pH であることが知られているが、pH 依存性の結合活性を呈するヒト抗体は、腫瘍特異的な集積を呈する格好の抗

体医薬品候補といえる。本研究開発によって、複数の興味深いヒト抗体およびそれらのアミノ酸改変抗体を見出すことが出来ており、今後の発展的な研究の展開に繋がる基盤的な成果が見出されている。

総括すると本研究開発計画は順調に進められたといえる。大量のヒト抗体アミノ酸配列情報を取得し、それらのヒト抗体クローンが生体内で経験した affinity maturation によるアミノ酸変化を正確に捉えることで、より高機能なヒト抗体の設計を目指したアミノ酸改変を目的とする計算科学的解析法の基盤開発を進めることができた。シングルセル・レパトア解析情報の計算科学的解析を基盤としたアミノ酸改変に基づく高機能抗体のデザインアルゴリズムの開発について、より一般的な任意のヒト抗体配列の高機能化を可能にするべく研究開発を継続していきたい。本研究開発計画では、そのような将来性・発展性が大いに期待できる基盤的なデータベースの構築と計算科学的技術の基盤開発という成果が得られたといえ、研究成果の今後のさらなる展開が期待できる。

The purpose of this research was to conduct detailed immune repertoire sequencing, including single cell analysis, focusing on B cell clones present in human disease tissues such as tumors, as well as human B cell clones that specifically recognize tumor antigens. Based on genome informatics of immune repertoire sequencing information, we aimed at developing a foundation model of informatics technology capable of estimating highly functional antibody clones, at the amino acid sequence level. In the development of antibody drugs, it is necessary to search for highly functional clones by step-by-step amino acid modifications; thus, both time and cost are major problems. It is highly demanded to develop informatics techniques for the efficient designing of amino acid modifications.

This research was conducted with four steps, obtaining the following achievements.

(1) Acquisition of comprehensive immune repertoire sequencing data

B cell clones that infiltrate into human clinical tissues and that recognize sulfated glycosaminoglycans, which were shown to be novel cancer antigens in our previous studies, were purified and single-cell immunoglobulin repertoire sequencing was performed. We succeeded in acquiring a large amount of sequence information of immunoglobulin clones. We could also establish our computational scientific approaches to profile the precise amino acid sequences as a combined dataset of heavy and light chain sequences per single cells.

(2) Development of design technology for highly functional antibodies using genome science

By computational science using the amino acid sequence information of above-mentioned immunoglobulin clones, we sought to develop bioinformatics pipelines to decipher characteristic amino acid sequences and their changes during affinity maturation, which are specifically related to cancers and antigens.

(3) Expression and purification of antibody molecules

Based on the sequence information of above-mentioned human immunoglobulins, we experimentally expressed and purified the actual antibody molecules. Their expression levels and functionality were evaluated to investigate the positive and negative effects of the designed amino acid modifications.

(4) Verification of biological functions of antibodies and feedback to machine learning

For the antibody molecules identified and designed in (1) to (3) above, we proceeded with their functional verification. In addition to the evaluation of expression stability, tumor-specific and sulfated glycosaminoglycan-specific antibodies were investigated for their binding abilities to various cancer cells and anti-tumor effects using mouse models. As a result, we have partially succeeded in identifying and designing multiple of highly functional antibodies that showed potentially intriguing functionalities.

In summary, this research project has been well proceeded as initially planned with intriguing achievements. The constructions of fundamental repertoire database and computational science technology were achieved that are expected to have promising future potentials.