

## 研究課題別事後評価結果

1. 研究課題名：染色体および RNA の機能変化からの疾患の系統的解析
2. 研究代表者名及び主たる研究参加者名(研究機関名・職名は研究参加期間終了時点):

研究代表者

油谷 浩幸 (東京大学先端科学技術研究センター 教授)

### 3. 研究実施概要

テーラーメイド医療の実現に向けたゲノム解析を推進する本研究領域の中で、本プロジェクトはヒトゲノムの多様性を探ることを中心課題に掲げた。同一の診断名がつけられた疾患であってもその発症機序が均一でなければ、症例毎に治療応答性は異なることになる。従って最適な治療を実現するためには症例を層別化することが必要となり、層別化を行うためのバイオマーカーとして個人のゲノムの違いに関するマーカーが求められる。プロジェクト開始の平成 16 年(2004 年)はようやく 10 万箇所程度の SNP を測定できるアレイが開発された時期であり、SNP と疾患のゲノムワイドの関連解析(GWAS)に世界中が着手し始めた時期であった。当時癌細胞ゲノムに多くの異常が生じることは周知であったものの、健常人のゲノムに SNP 以外にどの程度のゲノム多様性があるのか、さらには RNA が転写される際の多様性についてはほとんど解明されていなかった。

そこで個人のゲノムがどの程度異なるのかという疑問に答えるべく、染色体および RNA 変異の系統的解析技術開発、アレル間の遺伝子発現量の多様性という 3 つの研究項目を掲げ、最終的にそれらの多様性と疾患との関連づけを目指すという予定で研究を開始した。

#### (1) 染色体変異の系統的解析技術開発

##### ヒトゲノムコピー数解析

健常人ゲノム中に配列の一部が重複する *segmental duplication* という現象が存在し、時に疾患の原因ともなることが知られていたが、どの程度広汎な現象であるのかは全く不明であった。癌細胞ゲノムに生じる染色体変異の解析のため、すでに SNP アレイを用いたコピー数解析法 *Genome Imbalance Map* の開発に着手していたので、GWAS のために SNP アレイの高密度化への開発が急速に進んだことは幸運であった。今でこそゲノム全域をカバーするタイリングアレイが存在するが、当時は 50 万箇所の SNP を一度に測定できるアレイが最も網羅的な解析プラットフォームであった。トロント子供病院、Sanger センター、ハーバード大学、Affymetrix の研究者らとヒトゲノム構造多様性解析コンソーシウムを結成することとなり、HapMap プロジェクトに用いられた白人、アジア人、アフリカ人合計 270 検体について、本研究チームは専ら SNP アレイを用いた解析を担当することになったため、新たにコピー数推定アルゴリズム GEMCA (*Genotyping Microarray based CNV Analysis*) を開発した。

BAC アレイ解析結果と統合した結果、ヒトゲノム中に少なくとも 1447 箇所の CNV (*copy number variation*) が存在し、CNV 領域は合わせると 360Mb (ヒトゲノムの 12%) にわたり、3000 近い遺伝子が含まれるなど、ヒトゲノムにはコピー数多型が従来の予測以上に高頻度に存在し、その頻度は人種により異なることを明らかにした (*Nature* 2006)。その後さらに高密度のコピー数解析専用のアレイを用いて CNV のアレルと周囲の SNP との連鎖不均衡の関係の詳細な解析を行った結果、SNP によって tag 可能な CNV は全体でわずかであると考えられた。高密度アレイは Gorlin 症候群および全前脳症の染色体変異の解析にも有用であり、将来的に臨床診断にも用いられると期待される。

##### 腫瘍組織における染色体変異

癌において生じたゲノム変異を明らかにすることは、今後分子標的医薬品の開発が進むことにより、「癌のテーラーメイド医療」においてますます重要になると考えられる。SNP アレイを用いて癌細胞ゲノムにおける染色体変異について、肝細胞癌、子宮体癌、グリオーマ、肺癌、大腸癌、卵巣癌などのアレル別コピー数解析を行った。微細な染色体変異を検出するためには同一患者由来の正常 DNA を対照検体として用いることが重要である。SNP 情報を統合することによりアレル別にコピー数解析を行う利点として、片側アレルが失われ、もう一方

のアレルが増幅した **uniparental polysomy** を正確に検出できた。ホモ欠失領域の検出は通常のアレイ CGH では困難であるが、SNP アレイを用いた解析ではアレル毎にコピー数を計測することにより1アレル分の変化に相当するシグナルを推定できるので、臨床検体の解析においても LOH のみならずホモ欠失の鋭敏な検出が可能となる。

肝細胞癌でのホモ欠失が認められた **CSMD1** 遺伝子は大腸癌で高頻度に配列異常が報告されており、染色体変異、DNA メチル化など様々な機序により遺伝子の不活化が生じるものと推定され、今後癌細胞ゲノムに生じたゲノム変異の意義を評価する上でも考慮する必要があると思われた。

悪性膠芽腫では新規癌抑制遺伝子候補を見出した。この遺伝子はグリア細胞の発生・分化に重要な役割を果たすと示唆された。子宮体癌の解析ではコピー数変異とマイクロサテライト不安定性の関係について統合的に解析から明らかにした。

## (2) アレル間の遺伝子発現量の多様性解析

### アレル間遺伝子発現量解析

同一の転写制御下にありながら2つの染色体からの遺伝子発現量が異なる現象は、染色体の機能の違いを検出するために極めて有効である。SNP を利用して2つの染色体からの遺伝子発現量比較をゲノムワイドに行う方法 **Expressgenotyping 法 (EG 法)**<sup>TM</sup>を開発した。アレル間の発現量比の検証は SNP 部位を用いて cDNA をテンプレートに作成した **primer extension** 産物を質量分析することにより行った。ゲノムワイドな片側染色体に偏る発現を示す遺伝子の検出により、X 染色体の片側アレルの不活化や既知のインプリンティング遺伝子が確認された。

薬剤応答性について薬剤投与症例において遺伝子発現応答を調べることにより応答性に関わるバイオマーカーを見出そうとする試みはよく行われるが、臨床試験において多数症例を解析するには多大な時間および費用がかかる。EG 法によるアレル別遺伝子発現を調べることにより、前臨床の段階で薬剤応答性に発現多様性を示す遺伝子をヒト培養細胞でスクリーニングすることが可能である。不死化リンパ球株をホルボ・ルエステル (**phorbol myristate acetate, PMA**) および **ionomycin** で刺激 30 分後に EG 法により発現応答の異なるアレルのスクリーニングを行ったところ、測定された遺伝子の1%程度に3倍以上発現量の異なるアレルの存在が認められた。

薬剤刺激を行うことによって初めて検出されるアレル間の発現多様性が数多く存在する。薬剤応答性のみならず、増殖刺激やストレス応答における個人差のスクリーニングにも有用な解析技術になるものと期待される。

### 薬剤応答性の遺伝解析

薬の効果や副作用の現れ方には個人差があり、その機構を解明することは、医薬品の有効性と安全性の向上、臨床試験の効率化、および医療経済面からも重要な課題となっている。不死化リンパ球細胞株を用いた薬剤感受性測定モデルは、人への薬剤投与が不要なために解析対象とする薬剤が限定されず、多様な薬剤の感受性を測定できる系である。本研究はこのモデルを用い、ゲノムワイド連鎖解析などの遺伝子網羅的な解析手法によって薬剤感受性に関連する遺伝子の探索を行い、その手法の有効性を検討した。

不死化リンパ球細胞株は同じ操作によって不死化された細胞株であるので、測定された薬剤感受性の **variation** は遺伝子の個人差を反映していると考えられる。細胞株パネルを用いたハイスループット測定系を用いることで、多数のサンプルの薬剤感受性を繰り返し精度よく測定することが可能となった。エラーを除去した **informative SNP** を用いて算出した場合にはノイズが消え、**Identical by decent (IBD)**が 0, 1, 2 のいずれかに確定できた。一方、マイクロサテライトマーカーのタイピングデータを用いて IBD を算出した場合は、情報量が少ないために IBD が確率的にしか求まらない領域や、狭い領域では IBD の変化を検出できない場合があった。

本研究で測定した 3 種の抗癌剤 (5-フルオロウラシル、パクリタキセル、およびカンプトテシン) はそれぞれ作用機序の異なる薬剤であるが、これらの3薬剤同士の感受性に連鎖する染色体領域が認められ、互いに相関が認められたことは感受性決定因子のなかに共通する因子が存在することを示唆すると思われる。

## (3) RNA 変異の系統的解析技術開発

近年、ゲノムから読み出される RNA には従来知られている mRNA、tRNA、rRNA 以外に **non-coding**

RNA、とりわけ microRNA などの small RNA の存在が知られるようになり、染色体機能の制御に重要な働きをしていることが示唆されている。140 万種類以上のエクソン候補配列に対して 4 種類のオリゴプローブからなるプローブセットがデザインされた Exon 1.0 ST アレイ (Affymetrix) を用いて転写産物の多様性の解析を行った。正常組織 88 検体に加えて癌細胞株 139、臨床腫瘍組織 151 からの RNA 検体についてのエクソンアレイ情報をデータベース化した。

次世代シーケンサーが実用化され、転写産物の多様性を解析するために RNA の配列解析を行うことにより、頻度情報のみならず、転写開始点、スプライシングバリエーションなど遙かに詳細な情報が得られるようになったことからシーケンス解析へと移行した。

#### 4. 事後評価結果

##### 4-1. 研究の達成状況及び得られた研究成果 (論文・口頭発表等の外部発表、特許の取得状況等を含む)

本研究ではヒトゲノムの多様性の探索を中心課題とした。テーラーメイド医療の実現には症例の層別化が必要であり、個人のゲノムの違いに関するマーカーが必要となる。このため、個人のゲノムがどの程度異なるのかを明らかにすべく、染色体および RNA 変異の系統的解析技術開発、アレル間の遺伝子発現量の多様性解析という 3 つの研究項目を掲げ、最終的にはそれらの多様性と疾患との関連づけを目指した。

健常人ゲノム中では配列の一部が重複する現象が存在し、時には疾患の原因となることが知られていたが、系統的な解析は行われていなかった。このため、ヒトゲノム構造多様性コンソーシアムを結成し、ヒトゲノム中には少なくとも 1447 箇所の CNV が存在し、この領域には約 3000 の遺伝子が含まれることを明らかにした (Nature 2006)。本研究項目に用いた SNP アレイによる解析は、腫瘍組織における染色体変異、先天的ゲノム異常の検出にも用いられ、その有用性が明らかにされた。次いで、二つの染色体からの遺伝子発現量の違いを SNP を利用してゲノムワイドに比較する Expressgenotyping 法 (EG 法) を開発し、不死化リンパ球を用いたスクリーニングでは、測定した遺伝子の 1% 程度で 3 倍以上発現量の異なるアレルが存在することを確認した。薬剤の効果、副作用の個人差の解明につながる重要な成果である。RNA 変異の系統的解析にはエクソンアレイを用い、正常組織、癌細胞株、臨床腫瘍組織についての情報をデータベース化した。

これらの研究成果は、88 編の原著論文 (及び 50 件の口頭及びポスター発表) として発表されている。ヒトゲノムの多様性を明らかにする手法の開発とその疾患との関連づけを目指した研究は、テーラーメイド医療の実現に貢献する大きな成果と言える。中でも CNV の系統的解析は世界に先駆けた研究であり、特記すべき成果である。

##### 4-2. 研究成果の科学技術や社会へのインパクト、戦略目標への貢献

本研究では、テーラーメイド医療の実現に向けて、染色体および RNA 変異の系統的解析技術開発、アレル間の遺伝子発現量の多様性解析という 3 項目の技術開発を行い、ヒトゲノムの多様性を明らかにするとともに、各種疾患との関連づけを行った。

CNV の系統的解析では、コピー数多型が従来の予測以上に高頻度に存在し、その頻度が人種により異なることを明らかにしたが、この解析においては、研究チームが新たに開発したコピー数推定アルゴリズム GEMCA が重要な役割を果たした。GEMCA は研究チームのホームページ上で公開されており、現在では世界中で利用されている。その後、コピー数解析専用的高密度アレイにより、更に多くの CNV を発見するとともに、従来法では検出不能であった微小な染色体異常を検出することを可能とし、先天性異常症の検討では、新たな発症機序の推定に結びつけた。また、腫瘍組織に生じた染色体変異の解析は、分子標的医薬品の開発が進めば、癌のテーラーメイド医療において重要な役割を果たすことが期待される。SNP を利用して二つの染色体からの遺伝子発現量比較をゲノムワイドに行う EG 法では、新たなインプリンティング領域が同定された。また EG 法は薬剤感受性における個人差の探索に有効な手法であり、薬剤応答の個人差のスクリーニング法など、既に実用化に向けた検討が始まっているのは、領域の目的にかなう成果として高く評価できる。RNA 変異の解析では、高密度化されたエクソンアレイを用い転写産物の多様性の解析を実施した。次世代シーケンサーを用いて行われる今後の検討の進展に注目したい。

#### 4-3. 総合的評価

本研究では、ヒトゲノムの多様性を解明するために、3つの研究課題を掲げ、技術開発と開発した技術のテーラーメイド医療への応用を目指した。ヒトゲノムにSNP以外にどのような多様性があるのかを明らかにすることは、テーラーメイド医療の実現には極めて重要な研究課題である。本研究は一貫してその課題に取り組み、CNV、遺伝子発現のアレル間多様性、及びRNA変異のゲノムワイドの高感度な解析技術を開発した。これらの技術開発から疾患との関連づけを含め多くの成果が得られているが、中でも世界に先駆けて行われたゲノムワイドのCNVマップの作成は特筆すべき成果である。また、遺伝子発現のアレル間多様性の解析技術が実用化に結びつきつつあることも、テーラーメイド医療の実現にとって重要な成果と考える。開発された各種の技術により、ヒトゲノムの多様性と各種の疾患の関連づけが更に進み、テーラーメイド医療の実現に大きく貢献してくれることを期待してやまない。