

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 5 月 31 日現在

機関番号：14301

研究種目：基盤研究（B）

研究期間：2009～2011

課題番号：21390176

研究課題名（和文）次世代シーケンサーを用いた薬剤代謝酵素の高精度ゲノム配列・遺伝子型決定法の確立

研究課題名（英文） The establishment of accurate sequencing and genotyping systems of genes encoding drug metabolizing enzyme and transporter by using next generation sequencing system.

研究代表者 関根 章博（SEKINE AKIHIRO）

京都大学・医学研究科・教授

研究者番号：30425631

研究成果の概要（和文）：

薬剤応答性の違いを調査するための重点領域である薬物代謝酵素やトランスポーターはファミリー内で高い相同性を有し、誤った遺伝子型決定や塩基配列が得やすく、正確なスクリーニング系を構築しなければ副作用や薬剤反応性の違いの原因となる遺伝子座同定や個別化医療実現の支障となる。そこで、他の領域と相同性を持つ 28 遺伝子に焦点をあて、次世代シーケンサーにて日本人 96 人の詳細多型地図を作成、正確な遺伝子型決定法を確立した。

研究成果の概要（英文）：

Our aim is to construct accurate screening system for pharmacokinetics(PK) to uncover genetic loci associate with drug responsiveness such as side effects or responder/non-responder. Genes related to PK well known to recognize high homology within superfamily, simultaneous amplification of different genomic regions in one test is causes of incorrect genotyped, incorrect sequence, and a fatal mistake in diagnosis. We identified all variations in 28 genes encoding drug metabolizing enzymes and drug transporters, CYP1A2, CYP2A6, CYP2C8, CYP2C18, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP3A4, OAT1, OAT2, OAT3, OCT1, OCT2, OCTN1, OCTN2, OATP1, OATP2, MATE1, MATE2, MRP1, MRP2, MDR1, PEPT1, PEPT2, PMAT, TPMT, and UGT1A1, recognized high homology in different regions, based on the combination the pair-end method of Next Generation Sequencer with an internally developed target specific long PCR system. In addition, we constructed a routine genotyping system of almost short variations (94%) based on combination Invader method and the long PCR. We planning to perform the routine screening in collaboration with other researchers focus on drug responsiveness.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009 年度	6,800,000	2,040,000	8,840,000
2010 年度	4,000,000	1,200,000	5,200,000
2011 年度	4,000,000	1,200,000	5,200,000
年度			
年度			
総計	14,800,000	4,440,000	19,240,000

研究分野：医歯薬学

科研費の分科・細目：境界領域・応用薬理学

キーワード：ファーマコゲノミクス、次世代シーケンサー、副作用、レスポonder/ノンレスポonder、チトクローム P450、トランスポーター、Pharmacokinetics

1. 研究開始当初の背景

疾患の原因や易罹患性に関連する遺伝子(多型)を同定する手段として Genome-Wide Association Study (GWAS) 用チップが成熟し、多くの多因子性疾患の易罹患性関連遺伝子座が相次いで同定されていた。GWAS は約 60 万(600K)多型を調べるチップから 1M チップに増大し、精度が向上していた。調査対象の多型 (tag SNP : 実際にタイピングを実施する SNP) から連鎖不平衡で関連遺伝子を同定することを原理とするゲノム広域をカバーする手法ではあったが、対象の多型は common variants が主であり、希少疾患や薬剤誘発の副作用の原因となる rare variants の検出は不十分な状況であった。特に、ファミリー内で相同性を有する薬剤代謝酵素やトランスポーター等の多型の検出や配列の個人差が大きい HLA 領域等については rare variants だけでなく、common variants の情報も正確に取得できない状況下にあった。(現在でも)。

一方、申請の1年程前(H20年)より次世代シーケンサー(NGS)が登場し、全ゲノム解析を主に common variants のみならず rare variants の同定をも可能にし、希少疾患や薬剤誘発副作用等の低頻度の表現型の原因を探索するツールとして期待された。ただ、申請時の NGS の性能はそのリード長、総リード数、精度ともに十分ではなく、コストも極めて高く当該研究(薬剤代謝酵素およびトランスポーターの正確な塩基配列決定)の遂行には大きな障害が予想されていた。NGS を用いて正確な配列情報を得るためには、理想的には該当領域を長くかつ深く(coverage を多く)リードすることが基本となり、予測される原因多型の遺伝子型頻度から少なくとも 50 例以上の症例の解析(常染色体で少ない方のアレル頻度が 1%程度の多型)ができるコストになることが必要であるためである。本研究遂行実現に大きな期待を持たせたのが 1) long read が実現できる次々世代シーケンサー(fragmentあたり 1000base 以上のリード)が H21 年市販予定であったこと、2)NGS の性能向上が H22 年度初頭(申請2年目)にその開発メーカー側で計画されていたことによる。しかし、前者は大幅に遅れ(実際は H23 年度)、後者は H22 年 3 月までずれ込んだ。これらの影響を受け、genotyping 手法は H21~22 年度に、NGS による当該領域のシーケンスは H22 年度に予備実験を、本試験は精度が向上し、コストが低減する H23 年度に変更することが余儀なくされたが、適宜計画を変更して本研究を遂行した経緯がある。当初の NGS による塩基配列決定は H22、23 年で各々 25DNA (計 50DNA) の予定であったが、計画変更と H22

年度までの NGS 実行準備体制を整備したことにより 96DNA に拡大できた。このことは、より低頻度の rare variants 検出に寄与する。しかし、一方では実務遂行が最終年度末までずれ込むことになるため公開の遅れが予想された。なお、当該領域の正確なシーケンス情報を取得するには通常的手法では現在も容易でない状況にあるので当該領域研究に従事する研究室には協力体制を可能な限り迅速に行う必要があることを考慮して進めている。

2. 研究の目的

薬剤応答性の違い(副作用発現、レスポナー/ノンレスポナー)に関連する遺伝子座を同定することは、投薬前診断にて個人に最適な薬剤を選択する上で極めて重要である。薬剤応答性関連座位を調査する際には 2 つの重点領域が存在し、1 つは薬剤の血中・組織内濃度に関わる Pharmacokinetics (PK) 遺伝子群で、残りは薬剤ターゲットとその調節に関わる Pharmacodynamics (PD) 遺伝子である。前者は幾つかの遺伝子の配列・多型情報を整備することで共通利用できるが、後者は各薬剤の特性によって異なるため薬剤毎で準備する必要がある。急展するゲノム解析技術は、ゲノム全広域から当該原因遺伝子座を容易に同定できるかのように予測されるが、ファミリー内で高い相同性を示す PK 遺伝子群は正確な調査が難しく、コピー数多型の存在でさらに解析が複雑となり、重点領域でありながら調査から脱落することになる。また、ゲノム配列は人種差が知られ、日本人 PK 遺伝子群情報を整備することは薬剤導出・導入に不可欠となる。そこで、日本人 PK 遺伝子群の詳細調査基盤を構築するため、ファミリー内で相同性を有する、あるいは高い相同性領域をゲノム内の他領域に持つ CYP1A2、CYP2A6、CYP2C8、CYP2C18、CYP2C9、CYP2C19、CYP2D6、CYP2E1、CYP3A4、OAT1、OAT2、OAT3、OCT1、OCT2、OCTN1、OCTN2、OATP1、OATP2、MATE1、MATE2、MRP1、MRP2、MDR1、PEPT1、PEPT2、PMAT、TPMT、UGT1A1 の①遺伝子型(Invader 法)決定法、②次世代シーケンサーによる配列(pair-end 法)の決定法確立を目指した。

3. 研究の方法

公開ゲノム・遺伝子情報(<http://www.ncbi.nlm.nih.gov/guide/genomes-maps/>)・多型情報(<http://www.ncbi.nlm.nih.gov/guide/variation/>、<http://www.1000genomes.org/>)から CYP1A2、CYP2A6、CYP2C8、CYP2C18、CYP2C9、CYP2C19、CYP2D6、CYP2E1、CYP3A4、OAT1、OAT2、OAT3、OCT1、OCT2、OCTN1、OCTN2、OATP1、OATP2、MATE1、MATE2、MRP1、MRP2、MDR1、PEPT1、PEPT2、PMAT、TPMT、UGT1A1

遺伝子の exon1 上流 10kb~最終 exon 下流 3kb の全領域をカバーするように long PCR (5~15kbp) プライマーをプログラムを駆使して設計した。プライマーは 1) 該当遺伝子と相同を示す領域を増幅させないように特異的に、2) 設計プライマー上に多型が存在しないように、3) 異なる PCR 産物で同領域を 2 回重複リードできるよう、4) さらに NGS によるシーケンス後にコピー数多型 (CNV) の影響を受けないようにプログラムを駆使してデザインした。各 PCR サンプルはバイオアナライザーと泳動ゲルバンドのサイズ測定から大きなゲノム構造の違い (structure variation) の存在を事前予測した。これを基質に、genotyping および NGS による塩基配列の決定を 96DNA で行った。

1) genotyping

Invader 法はサンガー法、TaqMan 法 Golden-gate 法と並んで精度の高い genotyping 法である。そこで、遺伝子型は前記 PCR 産物を基質に Invader 法にて領域内のほぼ全 SNP について決定した。同プローブは qPCR を実施し、ヘテロ接合性、CNV についての情報も取得した。また、Omni2.5M GWAS チップのタイピングを実施し、一部 (GWAS は相同性のため全多型の genotyping は結果が得られない) の SNP について結果を比較するとともに、CNV (コピー数多型) の領域を前者と併せて検証した。

2) NGS による塩基配列の決定

Long PCR 産物を covaris にて断片化の後、Illumina HiSeq1000 を用いて pair-end 法にて配列を決定した。この際、相同遺伝子間では正確な配列を得るために異なる Index で PCR 産物標識し、配列を決定した。なお、qPCR や配列決定からコピー数多型やヘテロ接合性の違いが予測された場合には Life Tech 社 Solid5500XL にて mate-pair 法による全ゲノム解析からゲノム構造の違いを同定した。

4. 研究成果

図 1 は、Cyp2C9 と他のゲノム領域との相同

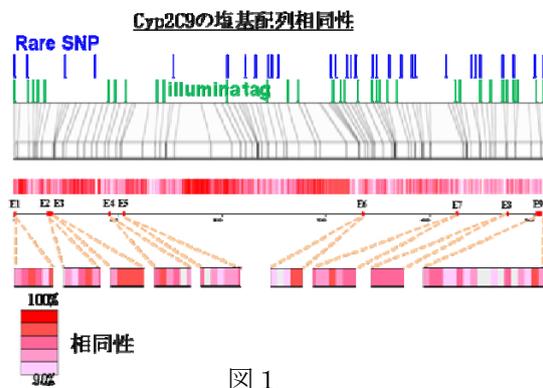


図 1

性を示す。上段、2、3 段目はそれぞれ rare variant (少ない方のアレル頻度=MAF<5%)、

Illumina GWAS Omni2.5M チップの tagSNP と common variants (MAF>=5%) を現わす。4 段目の他の相同性領域の存在をみても分かるように、これらの多型の大半は他領域と高い相同性を示す領域に位置する。5、6 段目は Cyp2C9 の exon 位置およびその相同性を示す。Omni2.5M チップは本図をみても分かるように高い相同性領域を避けてデザインされている。また、GWAS チップの特徴は common variant はかなりの検出ができる反面、rare variants のカバーは十分ではない。それ故、薬剤応答性の中でも副作用に関連する遺伝子座を確実に同定するツールとは言い難い。また、幸運にも GWAS によって副作用に関連する座位と連鎖不平衡にある tag が同定されても、その近傍に存在する副作用原因座位の genotyping は相同性のため正確な情報が得られない状況となる。

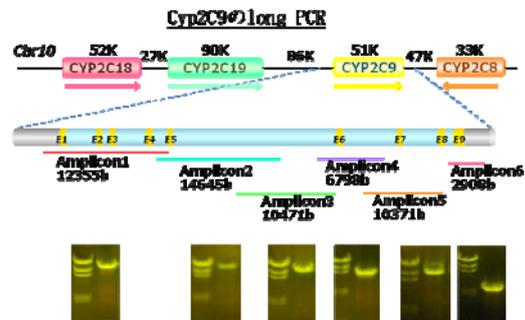


図 2

そこで、図 2 に示すようにリファレンス配列 (refSeq) を参照し、Cyp2C9 のみを特異的に long PCR 用のプライマーを独自プログラムを用いて設計した (2 段目)。また、プライマー上に多型が位置しないよう留意した。さらに、そのリスクを回避するために同じ領域を別の PCR 産物でカバーするようもう 1 セットの long PCR プライマーを設計した (図には未表示)。なお、設計された PCR 産物が目的の領域であることを事前に予測するため refSeq から PCR 長を算出し、ゲル (最下段) とバイオアナライザーにて検証を行った。

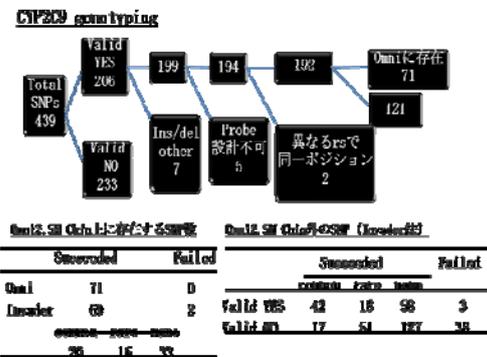


表 1

表 1 に示すように、Cyp2C9 には公開情報として 439 の short variation の報告がある。

うちタイピング実績のある多型 (valid YES) の 206 多型の 71 については Omini2.5M に搭載され、Invader 法によっても同一の結果が得られた(下段左表)。valid YES のうち、残る 135 多型のうち 118 多型については結果を得た。また、タイピング実績のない多型 (valid NO) についても同様にタイピングを行い 233 多型のうち 195 多型の結果が得られた(下段右表)。この結果は NGS によるシーケンス結果からも検証された。なお、下段右表に記載した多型については Omini2.5M チップでは設計できない領域でもあるので、独自の手法となる。また、NGS からは別の多型と思われる多型がいくつか同定されたが、サンガ一法による確認を行ったところ、一部を除きその大半が多型でないことを確認した。

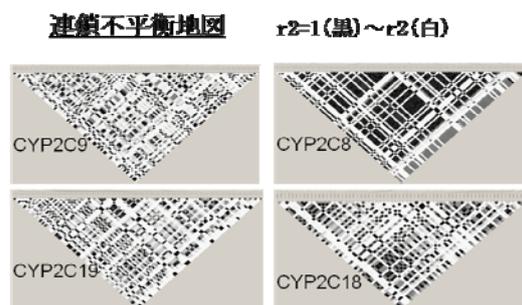


図 3

Cyp2C9 は特に Cyp2C19、Cyp2C8、Cyp2C18 との高い相同性を有する。各遺伝子上の多型が独立して genotyping できることから正確であることが示唆される。これらの多型の結果を用いて作成した連鎖不平衡地図が図 3 となる。同様の結果は他の PK 遺伝子でも得られている。

一方、Cyp2A6 上には 207 多型が報告されているが、Omni2.5M チップの tag は 5 多型で、うち 2 多型はタイピングミスを招く可能性があり、薬剤によってはその応答性調査で重要な遺伝子でありながらスクリーニングされない領域である。結果、これの多型については genotyping と sequencing により全て調査が終了し、うち 87% が genotyping できる状況になっている。また、Cyp2A6 には Structure valuation が知られ (図 4)、genotyping のみでは判定がつかない。NGS の結果から、下記に該当する Structure valuation (Cyp2A6*4) も 2 例同定された。

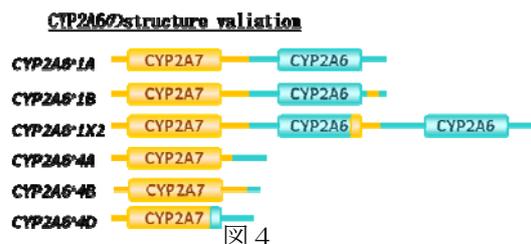


図 4

得られた全情報の報告は困難であるが、同様の結果は CYP1A2、CYP2D6、CYP2E1、CYP3A4、

OAT1、OAT2、OAT3、OCT1、OCT2、OCTN1、OCTN2、OATP1、OATP2、MATE1、MATE2、MRP1、MRP2、MDR1、PEPT1、PEPT2、PMAT、TPMT、UGT1A1 について該当する全領域のシーケンスを取得し、全 short variation の平均 94% の genotyping システムを完成した。これまで、PK 遺伝子群についてはこれまで十分なアプローチを遂行できなかった状況にあったが、本手法の確立により少なくとも薬物動態学的なスクリーニング法は確立したと確信している。特に、薬剤の導入導出時には最低限、今回完成した手法を試みることを望ましいと考えている。手軽に利用できるようにするために、現在、GWAS や NGS を開発した企業ならびに検査会社と協議を開始した。また、副作用 (ミオパチー、重症皮膚疾患、間質性肺炎、肝障害、腎障害、無顆粒球症など) を発現した検体を多く所持するいくつかの他施設にはすでに技術協力を行っており、96DNA に留まらず得られた情報の例数追加は今後も継続したいと考えている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 9 件)

1) Yoshimura K・Sekine A(3rd) : B-type Natriuretic peptide as an independent correlate of nocturnal voiding in Japanese women. *NeuroUrol Urodyn*, in press 2012 (PMID: 22532404) 査読あり

2) Hotta K・Mizusawa S(4th)・Sekine A (last) : Association between type 2 diabetes genetic susceptibility loci and visceral and subcutaneous fat area as determined by computed tomography. *J Hum Genet.* 57:305-310, 2012 (PMID: 22377712) 査読あり

3) Li H・Sekine A(13th) : Association of genetic variation in FTO with risk of obesity and type 2 diabetes with data from 96,551. East and South Asians. *Diabetologia* 55: 981-995, 2012 (PMID: 22109280) 査読あり

4) Hotta K・Mizusawa S (4th)・Sekine A (last): Genetic variations in the CYP17A1 and NT5C2 genes are associated with a reduction in visceral and subcutaneous fat areas in Japanese women *J Hum Genet* 57: 46-51, 2012 (PMID:22071413) 査読あり

- 5) Hotta K・Mizusawa S (4th)・Sekine A (last) : Computed tomography analysis of the association between the SH2B1 rs7498665 single-nucleotide polymorphism and visceral fat area. J Hum Genet 56: 716-719, 2011 (PMID:21796141) 査読あり
- 6) Hotta K・Mizusawa S (4th)・Sekine A (last) : Association of variations in the FTO, SCG3 and MTMR9 genes with metabolic syndrome in a Japanese population. J Hum Genet 56: 647-651, 2011 (PMID:21796137) 査読あり
- 7) Hotta K・Mizusawa S(5th)・Sekine A (last) : Association of the rs738409 polymorphism in PNPLA3 with liver damage and the development of nonalcoholic fatty liver disease. BMC Med Genet 11:172, 2010 (PMID :21176169) 査読あり
- 8) 関根章博・水澤精穂 (4th)・・・: 薬剤応答性の違いを調査するための Pharmacokinetics 関連遺伝子群の日本人における高精度遺伝子型およびゲノム配列決定法の確立 日本遺伝カウセリング学会誌 33 巻 1 号: 7-12, 2012 査読なし
- 9) 関根章博、前田士郎: 次世代シーケンサー Diabetes Journal 39: 143-147, 2011 査読なし

〔学会発表〕(計2件)

- 1) 関根章博・水澤精穂 (4th)・・・: 薬剤応答性の違いを調査するための Pharmacokinetics 関連遺伝子群の日本人における高精度遺伝子型およびゲノム配列決定法の確立 遺伝医学合同学術集会 2011、2011年6月11日(京都大学百周年時計台記念館)
- 2) 北本綾(他8名) 薬物代謝関連遺伝子群の高密度・高精度な遺伝子型決定法の確立 日本人類遺伝学会第56回大会 2011年11月10日(幕張メッセ)

〔図書〕(計1件)

関根章博 Medical Science Digest 特集 分子疫学コホートと新たな予防医学の展開 分子疫学研究とゲノム解析 484-487, 2011

〔産業財産権〕

ゲノム上に存在する多型や配列情報、その genotyping 手法、シーケンス手法は特許・知的財産権としては成立しない。日本人における薬剤応答性を調査するノウハウとなるので、特に日本の本領域研究者や海外からの薬剤導入・導出に関わる企業に役立てるプランを考えている。

〔その他〕

本研究の成果は、薬剤応答性試験を行う上での重点領域である PK 遺伝子群の詳細解析手技および基盤情報である。これを推進するには、①技術協力をを行う、②手法を開示する、③ツールとして多くの研究者が利用しやすいように一般化する、④配列の開示を行う等の方法が考えられる。①については、既に副作用を中心とした検体収集を行った研究室と討議し、技術協力を開始した。また学会発表などを通じて問合せがある場合には協力を行っている。②は当該領域を調査するためのプライマー情報、次世代シーケンス法、タイピングプローブ情報とその手法を開示する準備と論文化を進めている。③の稼働で広く活用可能となるので、タイピング・次世代シーケンサー開発企業、検査会社と討議し、多くの研究者が実行できるシステム開発を行うための話し合いを行っている。④も実現できるように準備を進めているが、ゲノム広域となるため個人情報・倫理面に留意し、開示方法などの体制を整備し、コントロールアクセスできるよう準備を進めている。

6. 研究組織

(1) 研究代表者

関根 章博 (SEKINE AKIHIRO)
京都大学・医学研究科・教授
研究者番号：30425631

(2) 研究分担者

水澤 精穂 (MIZUSAWA SEIHO)
京都大学・医学研究科・研究員
研究者番号：40523791