

戦略的創造研究推進事業 CREST
研究領域「テーラーメイド医療を目指した
ゲノム情報活用基盤技術」
研究課題「日仏共同体制による人種間
ゲノム多型の比較解析」

研究終了報告書

研究期間 平成15年10月～平成21年3月

研究代表者：松田 文彦
(京都大学大学院医学研究科、教授)

§1 研究実施の概要

研究の目的

ヒトゲノム計画の成果により、一塩基多型 (SNP) のゲノムワイド関連解析 (GWAS) を用いたヒトの複合遺伝性疾患の遺伝因子の同定が盛んに行われている。しかしながら解析には詳細な臨床情報をともなう多数の検体 (1,000 人規模) と高出力のタイピング設備に加え、最先端の統計遺伝学的手法の適用が必須である。本課題では、日仏国際共同研究で、解析対象疾患を免疫系疾患と一部の癌に限定して、白人と日本人で SNP を探索し比較することで、人種を超えて病気と関連する遺伝子・多型を同定し、疾患の原因遺伝子または疾患マーカーとなる SNP の組み合わせを発見することで、人種的偏差を加味した疾患別 SNP データベース構築を目的とする。また、さらに SNP と臨床情報を統合したデータベースの統計解析を行い、SNP に基づく遺伝学が治療に直結した次世代の臨床遺伝学として有効であることを実証し、疾患の分子レベルでの予知、診断および、将来の個人に合った治療や予後の予測 (テーラーメイド医療) のための情報の基盤を確立する。

研究実施内容

研究期間の前半は、候補遺伝子アプローチで、免疫関連遺伝子と DNA 修復関連遺伝子の SNP 同定から患者検体を用いた遺伝解析までを一貫して行なった。研究期間の後半は全ゲノムスキニングを取り入れ、免疫系疾患は関節リウマチ (RA) に限定して白人、日本人間で疾患感受性遺伝子を同定することとした。また、中間評価時に多数の検体収集の終了していた肺がんについても、同様の方法で感受性遺伝子検索を試みた。

1 SNP データベースの構築

研究室で生産される大量のジェノタイプング結果を格納し、また、解析結果に加え、検体の臨床情報、日々更新される公開データを遺伝解析を標準化して取り込むには、よく吟味されたデータモデルに基づいた優れたデータベースの構築が不可欠である。これらの点を念頭に置いて、リレーショナル・データベースシステム “MESHMD” を構築した。現在このデータベースの中には、本研究に関連するジェノタイプング結果 20 億データポイント以上が蓄積されている。また、一部のデータを一般公開した。さらに、GWAS で得られた日本人の健常者対照群の個人のジェノタイプ情報も、近々制限を設けた上で公開を予定しており、異なる疾患で GWAS を試みる研究者にとって有用なデータとなると考えている。

2 候補遺伝子解析

免疫系疾患と DNA 修復が関与する疾患を解析ターゲットにし、生物学的機能に基づいて選択した免疫関連遺伝子 169 個および、DNA 修復関連遺伝子 123 個について、フランス人 (白人) および日本人検体を用いて、遺伝子のエクソンと周辺領域およびプロモーター領域の塩基配列の再決定による SNP 同定を試みた。短期間で効率よく頻度の比較的高い SNP を同定するために、2 検体ずつを同じコピー数混合して、16 プールとし、免疫関連遺伝子からは 2,830 個、DNA 修復遺伝子からは 4,026 個の多型 (SNP、挿入・欠失) を同定した。そして、多人種におけるハプロタイプ構造の多様性に着目しながら、関連解析をより効率に行う SNP の同定および選別法を確立し、大規模ジェノタイプングに反映させた。

Illumina 社 Goldengate 法により、免疫関連遺伝子群から 1,536 個のマーカータグからなるパネルを構築し、日本人、ドイツ人の RA 患者、健常者検体 (日本人: 患者 238 例、対照群 184 例、ドイツ人: 患者 184 例、健常者 273 例) を用いたジェノタイプングを試みた。結果を人種ごとにプールして関連解析をおこない、さらにメタアナリシスの手法を用いて (DerSimonian-Laird 法) 多人種間での関連の強さを検討したところ、FCGR2A 遺伝子 (rs7551957: $DLp=8.58 \times 10^{-5}$)、DPYD 遺伝子 (rs6685859: $DLp=1.30 \times 10^{-4}$ 、rs7550959: $DLp=1.47 \times 10^{-4}$ 、rs7531138: $DLp=1.71 \times 10^{-4}$)、SPTB 遺伝子 (rs2269310: $DLp=7.89 \times 10^{-4}$) で、有意な関連が得られた。

3 RA のゲノムワイド比較解析

RA の GWAS による多人種間比較解析を、日本人、白人検体を用いて行った。まず日本人患者 674 例、対照群 934 例に対し、約 30 万 SNP に対してゲノムスキャンを行い、加えて約 55 万 SNP に対し、患者 317 例、対照群 298 例を追加解析した。その結果、*HLA* 遺伝子座に極めて強い疾患との相関が得られ ($p < 10^{-10}$)、用いた集団の有効性が示された。*HLA* 以外の領域に関しては、二度の解析で有意差が再現された SNP マーカーを順に選び、患者 933 例、対照群 855 例で再現性の検証をおこなった。

その結果、3 個のマーカーで再現性よく有意な結果が得られ、メタアナリシスの p 値もそれぞれ $MHp = 3.1 \times 10^{-6}$ 、 4.3×10^{-6} 、 3.4×10^{-5} が得られたため、現在最大の検体数で最終確認をおこなっている。

白人については、ドイツ人患者 192 例、対照群 839 例を用いて、30 万 SNP をスキャンした。日本人と同じく *HLA* 領域にきわめて強い関連が観察され、その他にも $p < 10^{-5}$ のマーカーが一個発見された。また、両人種の結果を用いてメタアナリシスを行ったところ、 $MHp < 10^{-5}$ のマーカーが 11 個得られ、そのうち一個は日本人で $p = 0.001$ 、ドイツ人で $p = 2.7 \times 10^{-4}$ と、有意差を保っており、現在検体数を増やして検証を行っている。今後は、候補領域のファインマッピングによる疾患感受性遺伝子/多型の同定、*HLA* やバイオマーカーを指標とするサブグループ解析、他のグループが施行した GWAS の結果との比較によるメタアナリシス等を行ない、疾患と関連の強いマーカーを同定後、RA 感受性遺伝子の同定を試みる。

4 肺がんのゲノムワイド比較解析

肺がんの GWAS による多人種比較解析を試みた。白人患者 1,989 例、対照群 2,625 例に対し Illumina 社 300K チップを用いて GWAS を施行し、染色体 15 番にきわめて強い関連を示すマーカーを 2 個 ($rs1051730 : p = 5 \times 10^{-9}$ 、 $rs8034191 : p = 9 \times 10^{-10}$) 検出した。 $rs8034191$ では、アレルの比較ではオッズ比 1.32 (CI : 1.21-1.45)、野生型ホモ接合体と変異型ホモ接合体のジェノタイプ比較では、1.80 (CI : 1.49-2.18) であった。またその結果は、患者 2,513 例、対照群 4,752 例を用いて、再現性が確認された。

これらのマーカーを含む 182kb の連鎖不平衡ブロック内には、3 個のニコチン様アセチルコリン受容体遺伝子 (*CHRNA3*、*CHRN4*、*CHRNA5*)、DNA 修復に関係する *PSMA4* 遺伝子に加え、*IREB2* (iron responsive element binding protein 2) と機能不明な遺伝子 LOC123688 が存在していた。現在のところ、疾患感受性多型を特定するには至っていないが、*CHRNA5* 遺伝子にアミノ酸置換を伴う変異 (D398N) が存在し、当該アミノ酸残基は種内、種間比較においても高度に保存されていることから、何らかの生理的役割を担っているのではないかと考えている。

また、サブグループ解析により、1) 腺がん、扁平上皮がん、小細胞がんの間には、大きなリスクの差はない、2) 喫煙者、非喫煙者の両方で有意なリスクの上昇がある、3) 頭頸部がんおよびニコチン依存症との関連はなさそうなことが明らかになった。また、当該領域は、日本人集団では肺がんとの関連は見いだされなかった。現在、日本人検体 (患者 719 例、対照群 1,494 例) を用いた GWAS を進めており、患者全例と対照群 980 例の解析が終了した。対照群の残りの検体の結果が判明し次第比較解析を試み、RA と同様の手法を用いた感受性遺伝子の同定を進める予定である。

5 研究成果の今後期待される効果

本課題が成功し、これらの疾患の感受性遺伝子の全貌を明らかにできれば、病態、人種における個々の感受性遺伝子の相対的貢献度が明確になり、疾患の多様な病態を分子レベルで説明することが可能となる。本研究が発展して、複数の遺伝因子と環境因子や生活習慣との関連が明らかにされると、1) 疾患の発症リスクの予測が可能で、予防医療に大きく貢献、2) 疾患のより客観的な診断 (分子診断) や予後予測が可能となることで、治療の方針決定を迅速におこなえる、3) 治療中の投薬にともなう有害事象の事前予測が可能となる、4) 疾患の発症の分子基盤の解明により、新薬の開発につながるなど、社会的イ

ンパクトもきわめて大きく、21 世紀の医療の中心となる「予防」に対する大きな貢献が期待され、「テーラーメイド医療」のよきモデルケースとなると考えている。

§2 研究構想及び実施体制

(1) 研究構想

ヒトゲノムの全容が 2003 年にほぼ解明され、ゲノム読解のみを追う時代は終わり、遺伝子多型からくる個人の遺伝子機能の変化と疾患との関わりに重きをおくようになってきた。ことにヒト複合遺伝性疾患は、複数の環境要因と遺伝要因が発症に関わる多因子疾患で、有病率が高い上発症期間も長く、多様な臨床病態を呈するため、医学的重要性に加え、行政的見地からも、国民の大きな関心事である。その遺伝因子が解明できれば、環境や生活習慣を変えることによる予防が可能となり、また個人にあった治療や予後予測（テーラーメイド医療）を可能にするため、21 世紀の医学・医療に大きな貢献が期待される。

複合遺伝性疾患に関わる遺伝子の多型として注目されているのは SNP である。SNP は高密度でゲノム全域に分布し、高頻度な多型であるため、全ゲノムにほぼ均等に存在すると考えられる複合遺伝性疾患の遺伝要因を探るマーカーとして優れている。また SNP は大多数が 2 アレルからなり、数値化及び解析が容易であることから、大量解析に向けたマーカーである。更に塩基配列決定やジェノタイプ技術の進歩により高速で大量かつ広範囲の解析が可能となるに伴い、遺伝解析の規模は個人レベルから集団レベルへと移った。これに呼応して、対象となる遺伝子多型も単一の多型から全ゲノムへ移行しており、世界各国で、多数の疾患に関する大規模なゲノムワイド関連解析 (GWAS) が行われている。

ヒトゲノムの多型は、人類の進化の過程での種としての多様性を大きく反映しており、人種間でその場所と頻度にかかなりの隔りがあることが知られている。ヒトゲノムの多様性は様々な病気と深い関わりを持ち、殊に複数の遺伝因子と環境因子の相互作用で発症する多因子型疾患の発症と進行の分子レベルでの理解には、疾患関連遺伝子・多型を同定し、その生物学的役割を解明することが必須である。しかしながら多因子型疾患は、病気に関連する個々の遺伝子の貢献度が低く、また人種によって各遺伝子の相対的貢献度に差があるため、単一人種のみを対象とした解析では遺伝因子の同定が困難である。特定の疾患に対し多人種の検体を用いた比較解析で、結果の再現性をもとに遺伝因子を同定することが最も有効な解析手法であるが、そういった積極的な国際協力体制はまだあまり多くない。

本研究課題のねらいは、京都大学とフランス国立ジェノタイプングセンターの日仏国際共同研究により、白人と日本人で SNP を探索し比較することで、人種を超えて病気と関連する遺伝子・多型を同定し、人種的偏差を加味した疾患別 SNP データベースを構築することにある (松田グループ、Lathrop グループ)。対象とする疾患を、免疫系の疾患と DNA 修復機構に関わる疾患 (おもに癌) に限定し、研究開始当初は候補遺伝子アプローチで、SNP 同定から患者検体を用いた遺伝解析までを一貫して行なったが、中間評価時の方針の再検討に基づいて、以前と比べてコストが飛躍的に安くなった全ゲノムスキニングを取り入れ、疾患を関節リウマチに限定して白人、日本人間で疾患感受性遺伝子を同定することとした。また、白人において多数の検体収集の終了していた肺がんについても、同様の方法で感受性遺伝子検索を行なった。

疾患感受性と関連する遺伝子・多型は、その生体での役割が解明されてはじめて生物学的意味を持つ。しかしながら、現在ヒト遺伝子の機能を細胞レベルで容易に判定することは甚だ困難である。そこで、SNP をもつ DNA 修復遺伝子の機能を、ニワトリ変異株を試験管がわりに用いたユニークな系をもちいて調べ評価する (園田グループ)。DNA 修復遺伝子の SNP のなかには抗癌剤のテーラーメイド治療を行う上で、薬剤の作用機構に影響を与える変異を含む可能性があり、既存の薬剤代謝に関する機構に依存する SNP データを相補する有用なデータになると考えられる。

また、絞られた候補 SNP 部位の臨床応用を図るための高精度ハイスループット SNP 解析

技術を確立するため、プライマー伸長反応を利用してコストを抑えかつ自動化、ハイスループットのタイピングプラットフォームの構築を試みた（辻本グループ）。

(2)実施体制

グループ名	研究代表者又は主たる共同研究者氏名	所属機関・部署・役職名	研究題目
松田グループ	松田 文彦	京都大学・大学院医学研究科・教授	人種間ゲノム多型情報の解析システムの構築
園田グループ	園田 英一朗	京都大学・大学院医学研究科・准教授	多型・変異をもつヒトDNA修復遺伝子機能の迅速同定法の開発
辻本グループ	辻本 豪三	京都大学・大学院薬学研究科・教授	高精度ハイスループットSNP解析技術によるSNP部位の確認
白川グループ	白川 太郎	京都大学・大学院医学研究科・教授	アレルギー疾患関連遺伝子項目の相互作用ネットワークの構築と機能解析
Lathrop グループ	Lathrop, Mark	Centre National de Genotypage (CNG), Centre d'Etude du Polymorphisme Humain (CEPH), Director	複合遺伝子疾患の多人種間の比較解析

§3 研究実施内容及び成果

3.1 サブテーマ1 人種間ゲノム多型情報の解析システムの構築と複合遺伝性疾患の多人種間の比較解析

(京都大学 松田グループ、フランス国立ジェノタイプングセンター Lathrop グループ)

(1) 研究実施内容及び成果

当サブテーマは、本研究課題の中核をなすもので、大規模ジェノタイプングプラットフォームを活用した、複合遺伝性疾患の感受性遺伝子の同定を目的とする。プロジェクト前半は、SNP データベースの構築と、免疫関連遺伝子、DNA 修復遺伝子の SNP 同定を通じた候補遺伝子解析を行ない、プロジェクト後半ではゲノムワイド相関解析 (GWAS) を用いた関節リウマチと肺がんの感受性遺伝子の探索を試みた。

1 SNP データベースの構築

大量のジェノタイプング結果を臨床情報とともに蓄積し、遺伝解析を行なえるフォーマットに標準化して提供するため、また、解析結果に加え、日々アップデートされる公開データも取り込むためには、よく吟味されたデータモデルに基づいた優れたデータベースの構

築が不可欠である。これらの点を念頭に置いて、以下の特徴・機能を備えたリレーショナル・データベースシステム MESHMD (Multiple Ethnic SNP Database for Human Multigenetic Disorders) を構築した (図 1)。

- 1) 複数のプロジェクトを管理する、ユーザー認証を導入したセキュリティーシステム
- 2) ことなる方法で得られるタイピング結果を、それぞれの解析結果ファイルより人手による加工なしに直接取り込むプログラム
- 3) 公共データベースから得られる情報と実験結果の標準化と融合
- 4) 得られた実験結果を用いた統計遺伝学的解析 (有意差の検定など) のプログラムの組み込み
- 5) 患者の臨床情報、人種などを指標とした基礎的な統計解析 (アレル頻度計算他) の自動化
- 6) 各種情報の確認と検索のための Web インターフェースによる検索ページの構築

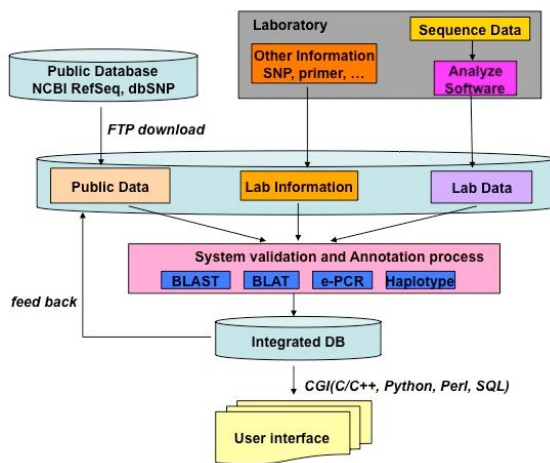


図 1 データベースのスキーム

データベースの構築は終了し、現在本研究に関連するジェノタイプタイピング結果 20 億データポイント以上が蓄積されている。データの一般公開を目指して、2006 年に試験的な公開を行なったが、セキュリティーに関する重大な問題点が発見され、内部用と、外部公開用の独立したデータベースを完全に分離して構築する必要が生じた。現在は外部公開用のデータベース用のハードウェアのセットアップが終了し、ユーザーを限定した試験の後、健常者における SNP の位置および頻度情報など、一部のデータの一般公開

を、<http://web.genome.med.kyoto-u.ac.jp/cgi-bin/snp/staff/top.cgi> で行なっている。

さらに、GWAS で得られた日本人の健常者対照群のジェノタイプタイピングの生データ (個人のジェノタイプ情報) も、近々制限を設けた上で公開を予定しており、異なる疾患で GWAS を試みる研究者にとって有用なデータとなると考えている。

2 候補遺伝子解析

1. 候補遺伝子の選定

本プロジェクトでは、自己免疫疾患をはじめとする免疫系の異常と、がんによって代表される DNA 修復が関与する疾患に限定して、ゲノム解析を行なった。そのため、これらの疾患に関係すると推定される既知の遺伝子、疾患の動物モデルの原因遺伝子のヒトのホモログをターゲットとして多型解析を試みた。

A) 免疫関連遺伝子 合計 191 遺伝子 (上記 URL にて公開)

- 1) ケモカイン、サイトカイン、それらの受容体、シグナル伝達に関与する分子など 151 遺伝子
- 2) 免疫系疾患の家系サンプルを用いたマイクロサテライト解析で疾患との関連が判明した染色体上の領域に存在する遺伝子 12 遺伝子
- 3) 免疫異常を示す動物モデルより推定される原因遺伝子座に対応するヒト遺伝子および、機能的にそれらの遺伝子と関わりを持つ遺伝子 28 遺伝子

これらのうち、性染色体上の遺伝子、多型が同定されなかった遺伝子、技術的な問題で多型解析が困難であった遺伝子をのぞく合計 169 遺伝子を、解析対象とした。

B) DNA 修復関連遺伝子 合計 123 遺伝子（上記 URL にて公開）

DNA の組換え、修復、複製、その他の DNA のメタボリズムに関する既知の遺伝子群（組換え修復、ヌクレオチド除去修復、塩基除去修復、ミスマッチ修復、チェックポイント関連、DNA ヘリカーゼ、DNA ポリメラーゼ、DNA スクレアーゼなど）

2. SNP の同定

上記の遺伝子群について、日仏の分担作業でフランス人（白人）と日本人の検体を用いて、遺伝子のエクソンと周辺領域およびプロモーター領域の塩基配列の再決定による SNP 同定を試みた。また DNA 修復遺伝子の約半数についてはコンゴ人（黒人）の検体も用いた。短期間で効率よく頻度の比較的高い SNP を同定するために、フランス人、日本人の健常者それぞれ 32 検体を、プロジェクト全体を通して SNP 解析の標準的パネルとして用いた（32 人の検体から 5% の頻度を持つ SNP を検出できる確率は、96% 以上であることが計算されている）。SNP 検索に用いる DNA 検体は、経費を抑えながら十分な情報が得られるように、2 検体ずつを同じコピー数混合して 16 プールとし、本研究で開発した SNP 検索ソフトウェアを用いて、4 本の染色体のピーク比をもとに 3 種のヘテロ接合体を鑑別したデータを、ハプロタイプ予想プログラムにかけてハプロタイプとその頻度を計算し、ジェノタピングに用いるマーカーを選択した（図 2）。

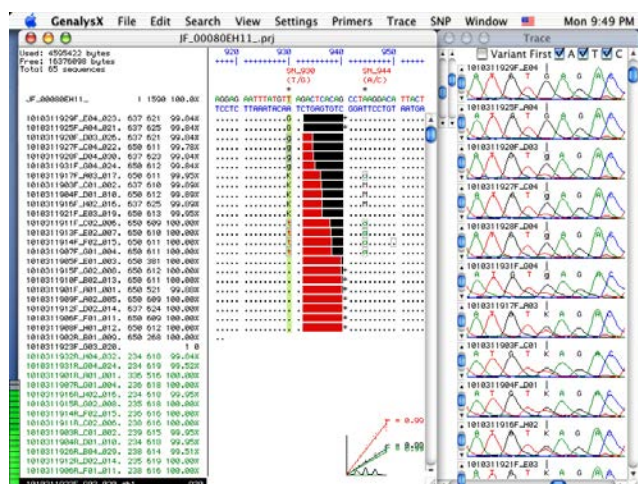


図 2 SNP 同定用ソフトウェア Genalys での解析例

- 1) いずれの人種でも、マイナーアレル頻度が小さくなるにつれ多型の検出数が増加し、多型の一般的傾向と合致した（図 3）。また、人種間共通 SNP および特異的 SNP の検討において、人種特異的 SNP はコンゴ人で最も多く見られた。また、2 人種間では、コンゴ人・フランス人共通の SNP が最も多かった（図 4）。
- 2) 遺伝子毎のハプロタイプ数は SNP 数に依存しない傾向が見られた。このことから、タグ SNP の選択で、多くの多型を持つ遺伝子も、限定された数の SNP によって解析可能であることを示していると考えられる。

同様の結果がフランス人、日本人の 2 人種で行った免疫関連遺伝子についても得られた。こういった知見をもとに、多人種におけるハプロタイプ構造の多様性に着目しながら、DNA 修復遺伝子群、免疫関連遺伝子群の関連解析をより効率的に行う SNP の同定および選別法を確立し、大規模ジェノタイピングに反映させた。

免疫関連遺伝子 169 個、DNA 修復遺伝子 123 個の SNP 同定をフランス人（白人）、日本人に加え DNA 修復遺伝子の半数についてはコンゴ人（黒人）で終了し、DNA 修復遺伝子では 4,026 個、免疫関連遺伝子では 2,830 個の遺伝的多型を同定した。そのうち DNA 修復遺伝子では 1,339 個（約 33%）が、また免疫関連遺伝子では 1,485 個（約 52%）が未記載の多型であった。その後、遺伝子ごと、人種ごとにハプロタイプを推定し、まず 3 人種で SNP 同定を行った DNA 修復遺伝子群において、人種間での多型の差を比較検討した。その結果、

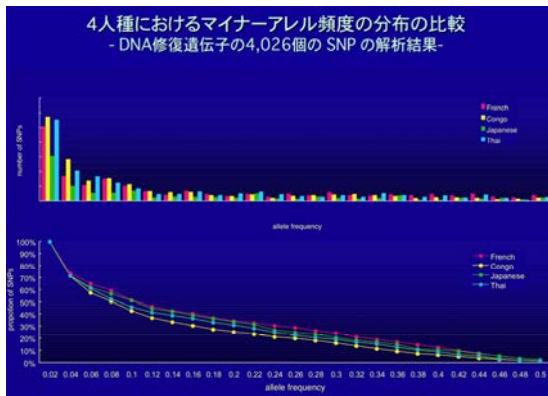


図3 DNA 修復遺伝子の SNP の人種間でのマイナーアレル頻度の分布の比較

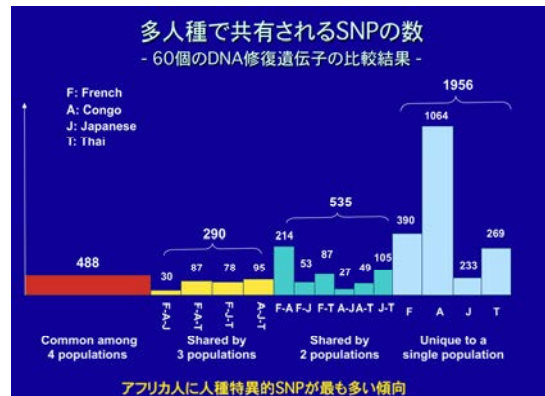


図4 人種間で共有される DNA 修復遺伝子の SNP の数

3. ジェノタイピングによる SNP と疾患の相関性解析

ハプロタイプ推定プログラムで、集団内で5%以上の頻度で出現すると予測されたハプロタイプについて、それぞれを識別可能な最少数の SNP をマーカーセットとして選択し、それらに HapMap プロジェクトのマーカーのうち本プロジェクトで同定されなかったマーカーを加え、検体数、SNP 数と密度に応じて、以下のテクノロジーの中から最適の方法を用いてタイピングを行った。

- 1) Goldengate 法：同定された多数の SNP を用いたマーカーパネルを構築し、網羅的にジェノタイピングを行う場合。
- 2) 塩基配列決定法：選択されたマーカーが小さな領域に集中して存在する場合
- 3) Taqman 法：特定の遺伝子を集中して解析する場合。また、Goldengate 法のタイピングの結果のバリデーションにも用いた。

各疾患に対して、人種毎にまず 96 例の患者と、同数の対照群 DNA 検体を用いたジェノタイピングを行った（一次スクリーニング）。有意差が得られたマーカーに関しては、疾患ごとにタイピングの結果の評価と、データの正確さを高めるための品質管理を施した上で、さらに検体数を増やして結果のバリデーションを行なう計画をたて、患者検体は日本人、白人ともに以下の複数の疾患で収集した。

- RA： 表2参照
- SLE： フランス人 150 家系（600 検体）、孤発例 100 検体、日本人 200 検体
- 喘息： フランス人 200 家系（2,000 検体）、英国人 4,000 検体、日本人 400 検体
- バセドー病： フランス人 150 家系（800 検体）、日本人 680 検体
- 肺癌： 表5参照
- 膀胱癌： フランス人 300 検体、日本人 260 検体
- 造血器腫瘍： 日本人 600 検体

これに加えて、健常者対照群の検体も収集した（表2、5参照）。

正常人と患者から得られた個々の SNP 及びハプロタイプを比較して統計処理し、疾患と関連を示す SNP または原因遺伝子座を直接同定することを目的に、以下のステップで結果の統計解析を行った。

- 1) アッセイ成功率の低い (85%以下) のマーカーは、統計解析から除外。
- 2) 各疾患の解析において、健常者対照群のジェノタイプのハーディー・ワインバーグ平衡 p 値が 0.05 を下回るマーカーは統計解析から除外した。
- 3) 各疾患の患者集団と健常者の間で、アレル頻度の差を検定した。

SNP 同定で得られた多型情報をもとに、Goldengate 法で大規模ジェノタイピングを試みた。両遺伝子群で、まず 1,536 個のタグ SNP からなるマーカーパネルを構築し、患者、健常者検体それぞれ 92 例を用いて一次スクリーニングを試み、タイピングデータの QC を終えた疾患から順に有意差の検定を SNP 毎、ハプロタイプ毎に順次行った。

複数の疾患のうち、RA の一次スクリーニングで人種を超えて潜在的有意差を示すマーカーが複数個得られたため、2006 年の中間評価において、最も検体数が多く興味深い結果が得られつつある関節リウマチに専念すべきとの助言があり、その後の解析は関節リウマチに限定して行なった。ここでは、そういった理由で、RA の解析結果についてのみ述べる。

一次スクリーニングと同じマーカーセットを用いて独立した検体セット (日本人患者、健常者それぞれ 146 例、92 例; 白人患者、健常者それぞれ 92 例、181 例) で再現性の検証をおこなった。結果を人種ごとにプールして関連解析をおこない、さらにメタアナリシスの手法を用いて (DerSimonian-Laird 法) 多人種間での関連の強さを検討した。その結果、メタアナリシスで $p < 10^{-2}$ を示す SNP が 22 個同定され、それらは 16 個の遺伝子に対応していた (表 1)。

Gene	dbSNP ID	Japanese freq. var.*		nominal Trend p	Caucasian freq. var.*		nominal Trend p	meta-analysis		
		RA	CTR		RA	CTR		DL p	OR (CI)	Dir.**
DPYD	rs6685859	0.023	0.050	0.0419	0.506	0.619	5.10E-04	1.30E-04	1.64 (1.27-2.12)	1
	rs7550959	0.166	0.215	0.0795	0.387	0.504	6.36E-04	1.47E-04	1.52 (1.22-1.89)	1
	rs7531138	0.168	0.215	0.0902	0.392	0.509	5.94E-04	1.71E-04	1.51 (1.22-1.88)	1
FCGR2A	rs7551957	0.098	0.170	0.0035	0.372	0.466	0.0062	8.58E-05	1.58 (1.25-1.99)	1
	rs1801274	0.176	0.225	0.0983	0.412	0.500	0.0109	0.0021	1.40 (1.13-1.74)	1
ITPKB	rs697846	0.202	0.283	0.0104	0.140	0.181	0.1391	0.0042	1.44 (1.12-1.85)	1
XPO1	rs1050567	0.236	0.331	0.0053	0.102	0.133	0.1658	0.0021	1.50 (1.15-1.94)	1
EIF2AK2	rs3770768	0.004	0.014	0.1516	0.099	0.155	0.0154	0.0082	1.72 (1.15-2.58)	1
CTLA4	rs926169	0.660	0.597	0.0909	0.423	0.357	0.0468	0.0087	1.31 (1.07-1.61)	-1
BCHE	rs2686399	0.844	0.819	0.3552	0.732	0.648	0.0064	0.0077	1.37 (1.08-1.73)	-1
ABCB1	rs6946119	0.121	0.159	0.1212	0.238	0.304	0.0276	0.0076	1.39 (1.09-1.78)	1
SPTB	rs2269310	0.368	0.475	0.0054	0.108	0.147	0.0731	7.89E-04	1.51 (1.18-1.92)	1
	rs229670	0.413	0.349	0.0613	0.268	0.177	0.0011	0.0024	1.48 (1.14-1.91)	-1
	rs2269304	0.118	0.075	0.0486	0.213	0.153	0.0185	0.0025	1.54 (1.16-2.05)	-1
IL4R	rs4787426	0.104	0.170	0.0117	0.122	0.153	0.2068	0.0076	1.49 (1.11-2.01)	1
ICAM2	rs4968681	0.323	0.409	0.0165	0.326	0.371	0.1656	0.0082	1.32 (1.07-1.62)	1
PECAM1	rs7503550	0.529	0.467	0.0960	0.566	0.474	0.0072	0.0018	1.37 (1.12-1.67)	-1
EVISL	rs537188	0.085	0.105	0.3384	0.140	0.212	0.0032	0.0045	1.52 (1.13-2.04)	1
VAV1	rs347033	0.144	0.227	0.0045	0.163	0.205	0.1124	0.0028	1.50 (1.15-1.97)	1
PLCG1	rs2866370	0.039	0.077	0.0217	0.050	0.078	0.0802	0.0054	1.81 (1.19-2.75)	1
TBX1	rs4819522	0.058	0.097	0.0419	0.182	0.254	0.0132	0.0014	1.58 (1.19-2.09)	1
	rs5746834	0.037	0.069	0.0390	0.190	0.248	0.0387	0.0049	1.52 (1.13-2.03)	1

表 1 RA の人種間比較解析で有意差が検出された SNP。*アレル頻度は NCBI build36.1 に対して変異となる塩基の頻度を表示。**Dir は、1 は変異アレルがリスクの増加に、-1 は抑制的にはたらくことを示す。

興味深いことに、FCGR2A 遺伝子の上流域に存在する SNP マーカー (rs7551957) は、日本人、白人ともに有意差を示し ($p=0.0035$ 、 $p=0.0062$)、かつメタアナリシスで $p=8.58 \times 10^{-5}$ という強い相関が得られ、またオッズ比も 1.58 (CI:1.25-1.99) であった (表 1)。FCGR2A 遺伝子は、IgG の Fc 部分との結合により抗体のエフェクター機能を担う分子群の一つで、マクロファージなどの貪食細胞に発現している。FCGR2A は SLE の感受性遺伝子としての

報告があり、また同じ群の *Fcgr3* 遺伝子欠損マウスは、コラーゲンで誘導される関節炎に対する耐性を示すこと、また *Fcgr2B* 遺伝子欠損マウスでは逆に関節炎が誘導されるなど、関節リウマチとの関係が強く示唆される分子であり、今後、機能解析を進めたいと考えている。

DPYD 遺伝子においては、白人において3個のマーカーで強い相関が得られ (rs6685859: $p=5.1 \times 10^{-4}$, rs7550959: $p=6.36 \times 10^{-4}$, rs7531138: $p=5.94 \times 10^{-4}$)、日本人においても rs6685859 で $p=0.0419$ と有意差が得られた (表1)。これらのマーカーは多型の頻度が人種間で大きく異なり、特に rs6685859 は日本人では2.3%の頻度と非常に低いが、白人では約50%の頻度を持ち、ほかの2マーカーについても17% : 39%と、白人における頻度をはるかに高いものであった。しかしながらリスクアレルは両人種で同じであり、*DPYD* 遺伝子の関節リウマチへの相対的貢献度は人種により異なると考えられる。*DPYD* 遺伝子はウラシルとチミジンの分解をつかさどる酵素で、この遺伝子が原因の常染色体劣性遺伝病 (デヒドロピリミジン尿症) が知られており、また血球系細胞では単球、リンパ球、顆粒球での発現が知られているが、その生物学的機能からは、自己免疫疾患との結びつきを示唆するデータは得られていない。

SPTB 遺伝子には、メタアナリシスで有意差を示すマーカーが3個同定されたが、各人種で最も強い相関を示すマーカーは異なっていた (日本人 rs2269310: $p=0.0054$ 、白人 rs229670: $p=0.0011$) (表1)。しかしながら、これらのマーカーのリスクアレルは人種間で一致しており、またマイナーアレル頻度が高い人種でより強い相関が見られることより、人種間で同一遺伝子において必ずしも同じ多型が疾患感受性に関与しているわけではないことを示すよい例であると考えている。*SPTB* 遺伝子は、赤芽球に特異的に発現するスペクトリンのベータサブユニットで、遺伝性球状赤血球症の原因遺伝子として報告されているが、自己免疫疾患との関連は報告例がない。

3 ゲノムワイド関連解析 (GWAS)

中間評価における方針変更で免疫系疾患は関節リウマチ (RA) に重点を置くこととなり、また全ゲノムのスキニングが予算的にも十分可能となったので、候補遺伝子を限定することなくゲノム全体をターゲットに、疾患感受性遺伝子の検索をおこなった。また、肺がんに関して、検体数が揃っていたので、同様の解析を試みた。

1. 関節リウマチ感受性遺伝子の同定

1) 検体収集

RAの診断は、アメリカリウマチ学会 (ACR) の診断基準項目に準拠しておこなった。それ以外にもリウマチの臨床研究に有用な臨床情報 (バイオマーカー、間質性肺炎などの合併症、重症度など) を検体採取時に収集した。そういった詳細な臨床情報を伴う日本人のRA患者の検体を、京都大学医学研究科免疫内科学教室が中心となって、国立相模原病院、道後温泉病院、東京大学アレルギー・リウマチ内科の協力で総数1,924検体収集し、また東京女子医科大学膠原病リウマチ痛風センターより1,264検体が供与され、合計3,188検体を用いた解析が可能となった。さらに研究協力機関から継続して検体収集をおこなっており、世界的にみても最大級のRA検体DNAバンクを構築することができた。また、健常者DNAに関しては、ファーマコ SNP コンソシアム (PSC) から約1,000検体の分与を受け、そのうち855検体を解析に用いた。さらに、愛知がんセンターの協力で、肺がんの遺伝解析のための健常者対照群1,494例を利用した。

白人検体に関しては、フライブルグ大学、モンペリエ大学より合計752検体収集し、さらにパリ大学を中心としたリウマチ研究コンソシアムとの連携で、検体収集が続けられており、また詳細な臨床情報も集積してきた。2008年度中に追加検体が1,000例以上集まる

予定で、バリデーションに用いる検体として十分数が得られた。

検体	RA 患者	例数	GWAS	検体	対照群	例数	GWAS
RA1	京都大学	674	300K	CTR1	JSNP	934	550K
RA2	京都大学	317	610K	CTR2	眼科系疾患	298	550K
RA3	京都大学	300	610K	CTR3	PSC	855	N.D.
RA4	京都大学	633	N.D.	CTR4	肺がん対照群	980	610K
RA5	東京女子医大	1264	N.D.	CTR5	肺がん対照群	514	N.D.
合 計		3188		合 計		3581	
RA11	フライブルグ大 学	192	300K	CTR11	フライブルグ大 学	210	300K
RA12	フライブルグ大 学	160	300K	CTR12	POPGEN	629	300K
RA13	モンペリエ大学	400	N.D.				
合 計		752		合 計		839	

表 2 RA の遺伝解析に用いた検体のまとめ。GWAS を行なった検体群は、チップ名を表示した。

2) 日本人および白人検体を用いた GWAS (30 万マーカー)

日本人 RA 患者検体 (RA1、674 例)、白人患者検体 (RA11、192 例) および白人健常者検体 (CTR11+CTR12、839 例) を用いて、Illumina 社 300K チップ (ゲノム上の 30 万個の多型マーカーを搭載) により全ゲノムスキャンをおこなった。日本人健常者対照群は、バイオバンクプロジェクトの公開データ (CTR1、934 例) を統計解析に用いることとした。

ゲノムスキャンを終了後、得られた結果の品質管理 (QC) を行なった。検体ごとの QC は、アッセイ成功率で行ない、95%以上の検体のみを用いた。また、マーカーごとの QC は、アッセイ成功率が 95%以上で、健常者におけるハーディー・ワインバーグ値が 0.05 以上のマーカーに限定して統計解析を試みた。

統計解析において、白人では各検体のジェノタイプが得られたので、通常に関連解析 (ジェノタイプ頻度、アレル頻度、優性、劣性の検定) に加え、用いた検体集団の構造化を補正するプログラムも用いた。日本人検体は、バイオバンクデータはジェノタイプ頻度のみの公開で、構造化補正が出来なかったため、通常に関連解析にとどめた。両人種において、HLA 遺伝子座に極めて強い疾患との相関が得られ ($p < 10^{-10}$)、用いた集団の有効性が示された。HLA 以外の領域に関しては、日本人検体で有意な p 値 (補正無しで $p < 10^{-5}$) を示す領域が 10 領域得られ、また 2 次スクリーニングの候補となりうる $p < 0.01$ を下回るマーカーは 2,400 個近く得られた。白人の解析は、日本人と比較して患者検体数が少ないため、結果の p 値が高めになっているが、補正なしで $p < 10^{-5}$ を示すマーカーが 1 個得られた。

3) 日本人検体の二次スクリーニング

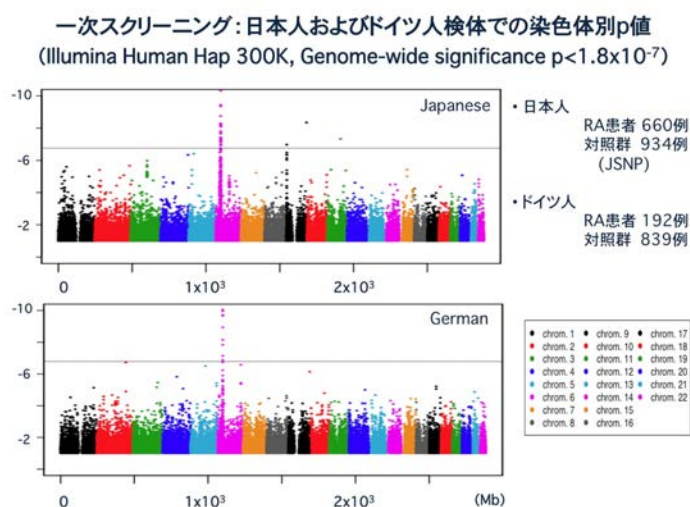


図5 縦軸は p 値の $-\log$ を表し、多重検定の補正の後も有意と判定できるボーダー ($p=1.8 \times 10^{-7}$) を水平の実線であらわした。

のマーカーを除外した後、有意差が再現性よく示された（一次 $p < 0.01$ 、二次 $p < 0.05$ ）マーカーから、メタアナリシス (Mantel-Haenszel 法) の p 値が小さいものよりまず 50 個選択し、更なる再現性確認の候補として選んだ。

4) 再現性の検証

各々のマーカーに対応するジェノタイピングプローブを Taqman 法でデザインし、独立したサンプルセット (患者: RA3+RA4、933 例、健常者: CTR3、855 例) で再現性を検証した。その結果、3 個のマーカーで有意な p 値が得られた (表 3)。K853227 を含む遺伝子は、単球、リンパ球、顆粒球での発現が知られているが、その生物学的機能からは、自己免疫疾患との結びつきを示唆するデータは得られていない。一方 K730060 を含む遺伝子は、抗体のエフェクター機能を担う白血球の細胞膜上の受容体で、炎症や自己免疫疾患との関連が示唆される分子をコードしていた。K965621 を含む遺伝子は、RA 患者に多く存在する自己抗体に対する抗原分子をコードしており、RA との関連が十分に予想される。現在これらのマーカーをさらなる検体 (患者: RA5、健常者: CTR4+CTR5) を用いてタイピング中である。

GWAS で用いた検体とは独立の日本人検体 317 例 (RA2) を、Illumina 社 610K チップ (ゲノム上の 55 万個の SNP マーカーに加え、6 万個の CNV プローブを搭載) でスキャンした。一次と同様の QC を施し、基準を満たした検体とマーカーのうち、一次スクリーニングの 300K チップに搭載されたマーカーを選別し、関連解析をおこなった。対照群には、眼科系疾患でゲノムスキャンをおこなった 298 例 (CTR2) を用いた。

一次スクリーニング、二次スクリーニングの結果を比較して、まず HLA 領域

SNPID		Genotypes (Cases)				Freq.1	Genotypes (Controls)				Freq.1	Trend p	Meta-analysis	
		00	01	11			00	01	11				MHp	Odds ratio (CI)
K853227	1st	55	269	347	0.282	104	418	412	0.335	0.0015	3.1E-06	0.80 (0.73-0.88)		
	2nd	30	127	158	0.297	32	153	122	0.353	0.0331				
	3rd	87	396	441	0.308	115	367	366	0.352	0.0065				
	total	172	793	947	0.297	251	938	900	0.345	6.6E-06				
K730060	1st	252	325	91	0.379	417	422	94	0.327	0.0018	3.4E-05	0.83 (0.75-0.91)		
	2nd	120	153	37	0.366	144	137	26	0.308	0.0303				
	3rd	377	432	113	0.357	384	377	86	0.324	0.0380				
	total	749	910	241	0.366	945	936	206	0.323	3.7E-05				
K965621	1st	211	307	141	0.447	344	442	148	0.395	0.0040				

2nd	98	154	79	0.471	120	132	46	0.376	0.0009		
3rd	308	417	193	0.437	298	404	134	0.402	0.0377		
total	617	878	413	0.447	762	978	328	0.395	5.3E-06	4.3E-06	1.24 (1.13-1.35)

表3 日本人のGWASで有意差を示したSNPのまとめ。スクリーニング毎のアレル頻度とp値、結果を総合したp値およびメタアナリシスのp値を示した。

5) 日本人検体を用いたGWAS (25万マーカー)

上述の二次スクリーニングで、日本人RA検体317例(RA2)を、Illumina社610Kチップを用いてゲノムスキニングを行ったが、このチップに含まれる25万マーカーは、300Kチップに含まれておらず、新たな解析を必要とする。これらの25万マーカーは、白人の解析に主眼をおいて選択された30万マーカーにない、黄色人種で頻度の高いマーカーが多数含まれており、日本人の解析にはきわめて重要である。しかしながら検体数が317と少なく、十分な検出力を得られないと判断し、さらに300例(RA3)のゲノムスキャンを610Kチップで試みた。300Kチップの解析と同様のQCを施し、基準を満たした検体とマーカーを選択した。RA検体(RA2+RA3、617例)と健常者対照群(CTR1+CTR2+CTR4、2212例)をそれぞれプールしておこなった予備的解析において、300Kの解析で有意差を示すK730060に対応する遺伝子のファミリーに属し、きわめて似通った機能を持つ遺伝子で、 $p=8 \times 10^{-5}$ 程度の関連が得られている。現在さらに詳細な解析を続けている。

6) メタアナリシス

300Kチップを用いた日本人、ドイツ人のタイピング結果を用いて、Mantel-Haenszel法を用いたメタアナリシスを行った。HLA領域のマーカーは除外しp値を計算したところ、メタアナリシスで $p < 10^{-5}$ のマーカーが11個得られ、そのうち10番目は、日本人でRAと強い関連を示すPAD4遺伝子であった。大多数のマーカーでは日本人、ドイツ人のいずれかで有意差がみられるのみであったが、K940910は日本人、ドイツ人の双方で有意差を保持しており(それぞれ $p=1.0 \times 10^{-3}$ 、 $p=2.7 \times 10^{-4}$)、現在検体数を増やして検証を行っている。

11個のマーカー中1個のみがドイツ人でより強い関連がみられた最も大きな理由として、GWASに用いた検体数が192例と少なかったことがあげられる。これは、192例のGWASがほぼ終了した時期に、ウエルカムトラスト・ケースコントロールコンソシアム(WTCCC)より英国人検体を用いたRAのGWASが報告され、ジェノタイプの生データが入手可能となったため、白人の解析を中断し、日本人に集中することにしたことによる(後述)。

SNP ID	Gene	Japanese (Freq. var.)		German (Freq. var.)		Japanese (Trend p)	German (Trend p)	Meta-analysis (MHP)
		Cases	Controls	Cases	Controls			
K800754	Gene A	0.254	0.314	0.747	0.791	1.2E-05	0.063	2.4E-06
K839194	Gene B	0.491	0.415	0.686	0.673	4.2E-07	0.639	2.5E-06
K727924	Gene C	0.130	0.090	0.310	0.258	2.1E-05	0.040	4.4E-06
K808581	Gene C	0.128	0.089	0.495	0.433	3.1E-05	0.028	5.3E-06
K940910	Gene D	0.426	0.474	0.297	0.397	1.0E-03	2.7E-04	7.0E-06
K942492	Gene C	0.125	0.087	0.310	0.258	4.3E-05	0.039	8.1E-06
K949453	Gene E	0.745	0.685	0.714	0.681	9.6E-06	0.211	8.3E-06
K955847	Gene F	0.696	0.631	0.625	0.597	5.4E-06	0.316	9.0E-06
K942887	Gene C	0.128	0.090	0.310	0.259	4.6E-05	0.042	9.4E-06
K938243	PADI4	0.480	0.415	0.699	0.667	1.4E-05	0.228	1.1E-05
K861680	Gene G	0.381	0.309	0.484	0.483	5.1E-07	0.953	1.2E-05

表4 二人種のGWASの結果を用いたメタアナリシスのまとめ。

7) 今後の研究計画

①RA 感受性遺伝子の同定

現在までに得られている RA と強い関連を示す SNP を含む染色体領域の LD ブロックを、GWAS を行った検体のジェノタイプ及び HapMap 計画の情報をもとに特定し、領域内の高密度多型地図を構築する。用いるテクノロジーは、マーカー数やその分布にも依存するが、Taqman 法もしくは塩基配列決定によって行う。得られた SNP の位置、頻度の情報をデータベースに蓄積し、多型間の連鎖不平衡、ハプロタイプ情報と集団における推定頻度を得て、領域の主要ハプロタイプを区別可能なマーカーを選択する。続いて、RA 患者および対照群の全症例を用いて、候補遺伝子領域のマーカーのジェノタイピングを実施し、連鎖不平衡地図ならびにハプロタイプ地図を作成し、相関解析、ハプロタイプ解析によって、RA 感受性と最も強い相関を示す多型を同定する。

ごく最近開発されたテクノロジーによるヒトゲノム全塩基配列決定計画が、日本を含む何カ国かの先進国で開始された。このプロジェクトは世界レベルでまとまった検体数を用いて全ゲノムを再決定し、情報を共有するもので、それによって得られた多型情報が得られるようになると、高密度多型地図の構築を独自に行う必要がなくなる可能性があり、解析の進捗が加速されると考えている。

②サブグループ解析

京都大が中心となって収集した日本人 RA 検体は、種々のバイオマーカー、重症度、合併症などの詳細な臨床情報をもなったものが多数存在する。また、東京女子医大の協力を得て得られた検体も、さらに詳細な臨床情報に加え、病態の経時的変化もデータベース化されている。バイオマーカーとしては、リウマチ因子 (RF)、抗環状シトルリン化ペプチド (CCP) 抗体、抗核抗体などで、特に白人において、RA 感受性遺伝子の相対的貢献度が抗 CCP 抗体の有無で異なることがごく最近報告された。そこで、同様の解析を日本人検体で行う。また、特に重症度の高い例 (ムチランス、骨強直など)、間質性肺炎を合併する例を重点的に収集したため、これらを指標としたサブグループ解析を通して、より細かい病態に応じた感受性遺伝子の同定も可能で、将来的には RA の発症にとどまらず、予後に関わる遺伝因子の同定も試みたい。

HLA 遺伝子に関しては、過去の精力的な HLA のタイピングにより、HLA-DRB1 が人種を超えて RA と強い遺伝的相関があることが知られている。HLA-DRB1 は、日本人集団では DRB1*0405 およびこれと共通するエピトープ (shared epitope ; SE) を持つ対立遺伝子がリスクアレルとして報告されている。しかしながら HLA 遺伝子座は、RA の遺伝因子のうち三割程度に過ぎず、リスクアレルを持たない RA 患者において、重要な未知の感受性遺伝子が発見される可能性は大きい。現在 GWAS を施行した 1,291 例のうち RA1 の 674 例はすでに DRB1 のタイピングを終了している。そこで、RA2、RA3 (合計 617 例) について、DRB1 のタイピングを行い、また、RA4 についても、順次タイピングをおこない、データを蓄積する。そして、SE (+) 群と SE (-) 群に分類し、両群間及び健常者との間で関連解析をおこなう。強い関連が得られた SNP に関しては、再現性検討の後、詳細な多型地図を構築し、感受性遺伝子を同定する。

③多人種間比較解析

RA は、HLA 領域以外に人種を超えた疾患感受性遺伝子が同定されていない。これは、単一人種の GWAS 解析では多重検定の補正に耐えた SNP を優先する解析戦略をとること、また RA に中程度に寄与する SNP を効率よく選別する手法が確立されていないことが原因している。こういった問題点を解消するには、大規模 GWAS の結果の人種間比較が最も効率的であるが、アレイにより搭載された SNP マーカーが異なるため、単純比較が困難で (WTCCC の解析は Affymetrix 社のアレイを用いており、Illumina 社のアレイと 10% 程度の共通マーカーしか持たない)、タイピングを施行していない SNP のアレル頻度を LD を考慮

して推定する複雑な解析手法 (imputation) が必要なため、そのような解析をおこなった例は RA では知られていない。研究代表者らは既に imputation を行うプログラムを利用して解析をおこなった経験があるので、WTCCC の結果に imputation を施し、データベース内で標準化作業を行ったのち、メタアナリシスを行ない、人種を超えて有意差を持つマーカーを検索することが可能である。そういった方法で関連の強いマーカーを同定後、高密度多型地図を構築し、RA 感受性遺伝子の同定を試みる。

④RA 感受性遺伝子／多型の機能解析

同定された RA 感受性遺伝子の疾患との関わりを明らかにするために、得られた疾患感受性遺伝子の SNP 情報を考慮した機能解析を試みる。SNP の生物学的機能評価は、当該 SNP の遺伝子上の位置で方法が異なる。そこで、遺伝子産物のアミノ酸配列に変化をもたらす場合は、RA 感受性遺伝子の野生型及び変異型 cDNA を細胞株に導入し、野生型・変異型タンパク質の細胞内での局在や発現レベルの違いを免疫染色法にて調べ、SNP が遺伝子機能に与える質的影響を明らかにする。遺伝子の発現調節モチーフ上に存在する場合は、レポーターアッセイを用いて、SNP が転写調節に及ぼす影響を検証する。

遺伝子の 3'非翻訳領域に存在する場合は、マイクロ RNA (miRNA) の関与を想定し、当該遺伝子の 3'UTR の SNP が存在する領域が既知の miRNA の結合部位である場合、野生型・変異型 cDNA を細胞へ導入し、miRNA に特異的に結合した mRNA を定量し比較することで、SNP が miRNA の結合に与える影響を評価する。

SNP がイントロン、遺伝子間領域に存在する場合は、SNP の生物学的機能の実験的評価はきわめて困難で、現在でも標準的な方法が確立されていない。そこで、バイオインフォマティクスを用いた転写因子結合部位のモチーフ検索、哺乳類のゲノム配列のアラインメントによる高度保存領域の決定を試る。さらに、RA 感受性多型が同一染色体上の離れた場所あるいは他の染色体上の遺伝子の発現を制御している場合も考えられるため、“Global Gene Expression Database”を用いた *in-silico* 解析を実施する。このデータベースは 400 人あまりの白人の GWAS 結果と、発現アレイを用いた遺伝子発現情報が統合されたもので、ゲノム上の約 41 万 SNP と 55,000 の転写産物の間で、各々の SNP で発現に変化が見られる遺伝子の情報がまとめられており、既にいくつかの報告で疾患関連遺伝子同定に力を発揮している。

2. 肺がんの感受性遺伝子の同定

肺がんは、プロジェクト当初より解析すべき疾患として、検体収集も順調に進行していた疾患である。開始 2 年間で収集されたフランス人検体約 600 検体に加え、中央ヨーロッパ 5 カ国の共同研究により総計 1989 例 (LC11) が解析に利用出来ることとなった。そこで、まず白人検体の GWAS を開始し、開始期間をずらせて日本人検体の GWAS も行なった。

1) 検体収集

日本人検体は、愛知県がんセンターの肺がん検体 (LC1、719 例)、同センターのホスピタルマッチの健常者検体 (CTR4、5 合計 1,494 例) に加え、京都大学呼吸器外科 (LC2、595 例)、岡山大学呼吸器外科 (LC3、228 例) より検体を収集した。対照群は、RA の解析に用いた CTR1、CTR2、CTR3 群も解析対象とした。

白人検体は、International Agency for Research on Cancer (IRAC) でヨーロッパの 8 カ国で収集された肺がん患者および、ホスピタルマッチの対照群 (LC11、1,989 例、CTR21、2,625 例) を全例用いて GWAS を開始することとし、再現性の検証他は 5 カ所から集められた肺がん患者およびホスピタルマッチの健常者 (患者 : LC12、13、14、15、16 合計 2,513 例、健常者 : CTR22、23、24、25、26 合計 4,752 例) で行なうこととした。

検体	肺がん患者	例数	GWAS	検体	健常者対照群	例数	GWAS
LC1	愛知県がんセンター	719	610K	CTR1	JSNP	934	550K
LC2	京都大学	595	N.D.	CTR2	眼科系疾患	298	550K
LC3	岡山大学	228	N.D.	CTR3	PSC	855	N.D.
				CTR4	肺がん対照群	980	610K
				CTR5	肺がん対照群	514	N.D.
	合計	1,542			合計	3,581	
LC11	IARC	1,989	300K	CTR21	IARC	2,625	300K
LC12	Tronto (カナダ)	330	N.D.	CTR22	Tronto (カナダ)	453	N.D.
LC13	EPIC (西ヨーロッパ)	781	N.D.	CTR23	EPIC (西ヨーロッパ)	1,578	N.D.
LC14	CARET (西ヨーロッパ)	764	N.D.	CTR24	CARET (西ヨーロッパ)	1,515	N.D.
LC15	Liverpool (英国)	403	N.D.	CTR25	Liverpool (英国)	814	N.D.
LC16	HUNT/Tromso (北欧)	235	N.D.	CTR26	HUNT/Tromso (北欧)	392	N.D.
	合計	4,502			合計	7,377	

表5 肺がんの遺伝解析に用いた検体のまとめ。GWASを行なった検体群は、チップ名を表示した。

2) 白人検体を用いた GWAS (30 万マーカー)

Illumina 社 300K チップを用いて、IARC の検体全例 (LC11、1,989 例、CTR21、2,625 例) に対し GWAS を施行した。結果の QC は RA の GWAS に用いたものと同様の基準を適用し、また EIGENSTRAT を用いた集団の構造化解析で、明らかにアジア人と思われる検体を除外し、最終的に患者 1,926 例、健常者 2,522 例で関連解析を試みた。その結果、染色体 15 番にゲノムワイド有意差 ($p=5 \times 10^{-7}$) をはるかに下回るマーカーが 2 個 ($rs1051730:p=5 \times 10^{-9}$ 、 $rs8034191:p=9 \times 10^{-10}$) 検出された。より強い p 値を検出した $rs8034191$ は、T/C 多型で、C がリスクアレルにあたるが、年齢、性別、国で補正した後のアレルの比較ではオッズ比 1.32 (CI: 1.21-1.45) が得られ、ジェノタイプ C/T 対 T/T、C/C 対 T/T の比較では、それぞれ 1.27 (CI: 1.11-1.44)、1.80 (CI: 1.49-2.18) であった。また、喫煙歴と本数を考慮した補正後も、それらの値に大きな変化は見られなかった。その他のマーカーでゲノムワイド有意差を下回るものはなかったが、 $p < 5 \times 10^{-5}$ のマーカーが 29 個検出された。

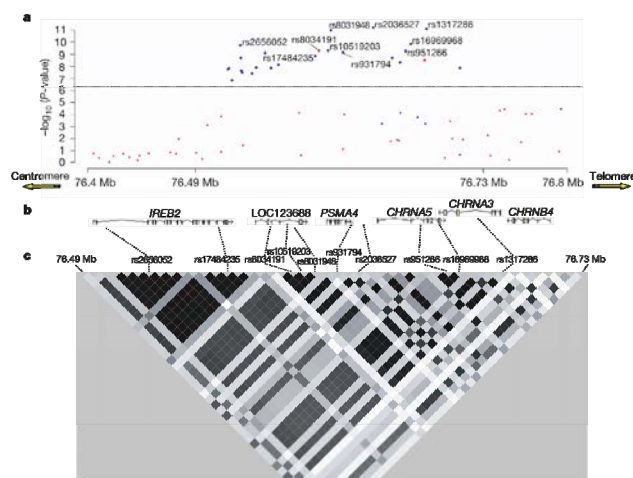


図6 肺がんと強い関連を示した染色体 15 番の領域。a. 一次スクリーニングのマーカー毎の p 値。b. 領域の遺伝子構成。c. 連鎖不平衡情報。

そこで、染色体 15 番の 2 個のマーカーに着目して、これらのマーカーを含む 182kb の LD ブロックを特定した (図 6-c)。きわめて興味深いことに、領域にはニコチン様アセチルコリン受容体遺伝子が 3 個 ($CHRNA3$ 、 $CHRN4$ 、 $CHRNA5$) に加え、DNA 修復に関する $PSMA4$ 遺伝子が含まれていた (図 6-b)。ブロック内の SNP から 300K チップに含まれていないマーカーのうち、上記の 2 マーカーと強い LD の関係にあるもの、領域内に位置する遺伝子上の SNP で、アミノ酸置換を伴うものなど 34 個を選

び、LC11、CTR21 全検体で Taqman 法によるファインマッピングを行った。結果は、34 個のうち 23 個のマーカーについて、ゲノムワイド有意差 ($p=5 \times 10^{-7}$) を下回る p 値が得られ、そのうち 2 個 (rs16969968、rs2036527) は、GWAS で得られた最も低い p 値を下回るものであった (図 6-a)。

さらに、有意差を示したマーカーのうち rs8034191、rs16969968 の 2 マーカーで独立した検体群をタイピングし、再現性の検証を試みた。検体群は、肺癌患者、健常者対照群ともすべての検体を用い (患者: LC12、13、14、15、16 合計 2,513 例、健常者対照群 CTR22、23、24、25、26 合計 4,752 例)、Taqman 法を試みた結果、独立した 5 検体群のすべてで有意差が再現され、6 検体群の総合で、rs8034191 において $p=5 \times 10^{-20}$ という、きわめて強い関連が得られた。アレルのオッズ比 1.30 (CI: 1.22-1.40)、ジェノタイプ C/T 対 T/T、C/C 対 T/T では、それぞれ 1.21 (CI: 1.11-1.31)、1.77 (CI: 1.58-2.00) が得られた。また、リスクアレルの頻度は 0.34 で人口における寄与危険度 (PAR) は 15% と計算された。

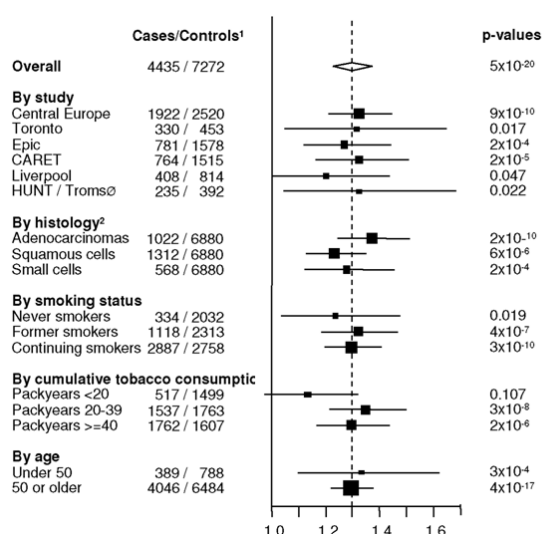


図 7 白人検体を用いた肺癌の GWAS の再現性検証とサブグループ解析

有意差は得られなかった。

3) 疾患感受性遺伝子の同定

疾患との強い関連がみられた領域中には、ニコチン様アセチルコリン受容体遺伝子サブユニット (*CHRNA3*、*CHRNA4*、*CHRNA5*)、DNA 修復関連遺伝子 *PSMA4* に加え、*IREB2* (iron responsive element binding protein 2) と機能不明な遺伝子 *LOC123688* の 6 個の遺伝子が存在していた。ファインマッピングでは、マーカー間の LD が非常に強いため、p 値から感受性多型を同定することは不可能であった。興味深いことに、*CHRNA5* 遺伝子の第 5 エクソンに位置する rs16969968 は、アミノ酸置換を伴う SNP の中で唯一有意差が得られ、受容体の第 2 細胞内ループの中心部にあたる 398 番目のアミノ酸をアスパラギン酸からアスパラギンへ置換 (D398N) するものであった。第 2 細胞内ループの生理学的機能は不明であるが、サブユニット間の種内比較、種間比較においてもアスパラギン酸が保存されていることから、何らかの生理的役割を担っているのではないかと考えている。しかしながら、領域内の他の遺伝子のイントロンおよび遺伝子間に位置する SNP の生物学的機能は不明であり、これらが真の疾患感受性多型である可能性は依然残っている。

4) 日本人検体を用いた当該領域の関連解析

当該 LD ブロックに含まれる SNP のうち、白人の解析で特に強い関連のあった領域から 16 個を選び、LC1 全例 (719 例) と健常者の一部 (CTR3 380 例、CTR4 785 例) を用いた Taqman 法によるジェノタイピングを試みたが、日本人集団において肺がんと有意な関連を示すマーカーは得られなかった (表 6)。当該領域は、白人において有意差が得られたマーカー間で $D' > 0.8$ 、 $r^2 > 0.6$ ときわめて強固な LD 構造がみられるが、日本人ではマーカー間の LD の低さが原因して、SNP 頻度にばらつきが多いためではないかと考えている。

Gene	SNP ID	nucleotide		Freq. var.	Caucasian		Japanese	
		Ref.	Var.		Trend <i>p</i>	Odds ratio (CI)	Freq. var.	Trend <i>p</i>
IREB2	rs2656052	A	C	0.33	2.51E-10	1.35 (1.23-1.49)	0.23	0.625
IREB2	rs7181486	T	C	0.33	3.16E-08	1.3 (1.18-1.42)	0.16	0.268
	rs931794	G	A	0.66	1.26E-08	0.77 (0.70-0.84)	0.63	0.073
CHRNA3	K1304830	---	CTCT		N.D.		0.88	0.19
CHRNA3	K1304829	---	CCCC		N.D.		0.51	0.061
CHRNA3	rs578776	G	A	0.27	2.51E-04	0.83 (0.75-0.92)	0.85	0.087
CHRNA3	K1304827	G	A		N.D.		0.28	0.138
CHRNA3	rs4887069	A	G		N.D.		0.78	0.833
CHRNA3	rs3743076	T	A		N.D.		0.28	0.281
CHRNA3	K1304816	C	G		N.D.		0.07	0.208
CHRNA3	rs3743075	T	C		N.D.		0.59	0.065
CHRNA3	rs3743073	G	T		N.D.		0.58	0.078
CHRNA3	rs8040868	T	C		N.D.		0.37	0.389
CHRNA3	rs1948	A	G	0.65	0.019	1.11 (1.02-1.22)	0.59	0.175
CHRNA4	rs28534575	T	G		N.D.		0.49	0.107
CHRNA4	K1304828	A	C		N.D.		0.29	0.085

表 6 染色体 15 番の候補領域の、二人種間での比較解析

5) 日本人検体の GWAS

illumina610K チップを用いた GWAS を、LC1 (719 例) とホスピタルマッチの対照群である CTR4 (980 例) を用いて施行した。QC を RA の解析と同等の条件で行い、関連解析を現在試みている。

6) 今後の研究計画

①日本人における肺がん感受性遺伝子/多型の同定と機能解析

RA の解析と同様の手法を用いて行う。まず、強い関連が得られたマーカーから順次再現性の検証を行う。検体は患者検体群 LC2、LC3 を用い、健常者は CTR3、CTR5 を用いる。また、既に GWAS を終えている CTR2、公開データの CTR1 も適宜参照しながら、候補遺伝子・多型を含む染色体領域を絞り込む。その後、領域の詳細な多型地図を構築し、感受性遺伝子を特定する。機能解析についても、基本的には、RA で用いる手法を踏襲する。

②多人種間比較解析

肺がんの解析は、両人種の解析を illumina 社のチップを用いて行ったため、RA と比べて比較解析がはるかに容易である。しかしながら、白人は 300K チップを用いたため、解析マーカー数が日本人の約 50% である。そこで、半数のマーカーについては、白人の結果に imputation を施した後に比較する必要がある。Imputation の結果をデータベース内で標準化

作業を行ったのちメタアナリシスの手法を行ない、人種を超えて有意差を持つマーカーの検索を試みる。

③サブグループ解析

がんのタイプ、予後等の臨床情報、喫煙歴をはじめとする生活習慣情報により、集団をサブグループ化し、白人と同様の解析を試み、比較検討する。

(2) 研究成果の今後期待される効果

関節リウマチ (RA) は、発症の機序が不明で予後予測も難しく、また発病期間が長く、治療も難しいため、患者にきわめて大きな肉体的・精神的負担を強いることなど、原因の究明と最善の治療の開発が緊要な課題である。

肺がんは、今や世界的にみてもがん死の1位を占め、その数は増加の一途をたどっている。発見・診断も困難で、またいわゆる「難治がん」と考えられ、治癒率も他のがんに比べ低いのが現状である。

本研究は、RA と肺がんの発症にかかわる遺伝因子を GWAS により網羅的に探索するもので、日本人で本課題と同規模 (数千人の検体) の詳細で質の高い臨床情報を伴う検体を用いた遺伝解析は聞かない。さらに、日本人と白人の大規模 GWAS の比較解析という、現在まで試みられたことのない独創的な手法を用いて、人種を超えた感受性遺伝子が同定できる可能性が大きい。本課題が成功し、これらの疾患の感受性遺伝子の全貌を明らかにできれば、病態、人種における個々の遺伝子の相対的貢献度が明確になり、また感受性遺伝子の組み合わせで、疾患の多様な病態を分子レベルで説明することが可能となる。本研究が発展して、複数の遺伝因子と環境因子や生活習慣との関連が明らかにされると、このような情報基盤を整備することにより、将来的には

- 1) 疾患の発症リスクの予測が可能で、予防医療に大きく貢献
- 2) 疾患のより客観的な診断 (分子診断) が可能となる
- 3) 予後予測が可能となることで、治療の方針決定を迅速におこなえる
- 4) 治療中の投薬にともなう有害事象の事前予測が可能となる
- 5) 疾患の発症の分子基盤の解明により、新薬の開発につながる

など、社会的インパクトもきわめて大きく、21 世紀の医療の中心となる「予防」に対する大きな貢献が期待され、「テーラーメイド医療」のよきモデルケースとなる。

3.2 サブテーマ2 多型・変異をもつヒト DNA 修復遺伝子機能の迅速同定法の開発 (京都大学 園田グループ)

(1) 研究実施内容及び成果

研究の目的

疾患に関連する SNP が遺伝子機能に与える生物学的影響について、SNP をもつ DNA 修復遺伝子の機能を細胞レベルで容易に判定できるというユニークな系をもちいて調べる。具体的には、DNA 修復遺伝子を欠損したニワトリ細胞株のコレクションにヒト DNA 修復遺伝子の野性型及び変異型の cDNA を導入し放射線や薬剤に対する感受性を測定することにより、同定した多型の機能を迅速に評価する。DNA 修復遺伝子の SNP のなかには抗癌剤のオーダーメイド治療を行う上で、薬剤の作用機構 (特に DNA を直接傷つけるもの) に影響を与える変異を含む可能性があり、既存の薬剤代謝に関する機構に依存する SNP データを相補する有用なデータになると考えられる。

研究実施方法

現在までに約 30 種類の DNA 修復欠損 DT40 細胞株が樹立されており、ヒト遺伝子がこれ

らのニワトリ細胞株の機能欠損を正常に相補することを利用して、エクソンの翻訳領域の SNP の生物学的意義を解明する。

1. SNP 同定に用いた検体のなかで、翻訳領域にアミノ酸置換をひきおこす SNP を持つ検体と持たない検体の細胞株より mRNA を単離し、全長 cDNA を合成する。
2. 発現ベクターに組み込んだ cDNA を当該遺伝子の欠損した DT40 細胞株に導入し、遺伝子導入細胞株の、薬剤、放射線、紫外線などの外来変異源に対する感受性を調べることにより、SNP がタンパクの機能の変化にどの程度関与しているかを明らかにする。
3. 得られた結果を、現在進行中の薬剤に対する SNP データ（多くは薬剤代謝と相関）とともに統計処理して、薬効と副作用にどの程度のインパクトを持つかを検討する。

成果

DNA 修復遺伝子の中で、相同組換えに関与する *XRCC2*, *XRCC3*, *RAD51* の白人にみられる SNP を選び、発現ベクター(CMVpromoter-ires-GFP)にクローニングし、それぞれ *XRCC2*^{-/-}, *XRCC3*^{-/-}, *RAD51*^{-/-} DT40 細胞に導入した。まず GFP の発現を指標に発現陽性株をスクリーニングしたのち、RT-PCR または Western blot 法により、野生型と同程度の遺伝子または蛋白発現株を複数樹立し、細胞増殖を液体培地で、また抗癌剤シスプラチンに対する感受性を半固形培地中で測定し、野生型のそれと比較した。

本研究で調べた 3 種類の相同組換えに関する遺伝子のうち、*RAD51* および *XRCC3* の SNP は、細胞増殖・抗癌剤感受性（シスプラチン）に関して野生型とほぼ同様な機能を示した。興味深いことに、*XRCC2* の SNP (188His) は正常な機能を回復するのみならず、シスプラチンに対して野生型よりさらに耐性を示した（図 8）。

遺伝子	変異	頻度(%)	増殖能	シスプラチン感受性
<i>XRCC2</i>	188Arg to 188His	8.6	野生型と同じ	野生型より耐性
<i>XRCC3</i>	241Thr to 241Met	31.7	野生型と同じ	野生型と同じ
<i>RAD51</i>	56Pro to 56Ser	1.7	野生型と同じ	野生型と同じ
	5'UT G to T	22.5	野生型と同じ	野生型と同じ

表 7 DT40 細胞を用いた機能評価実験のまとめ

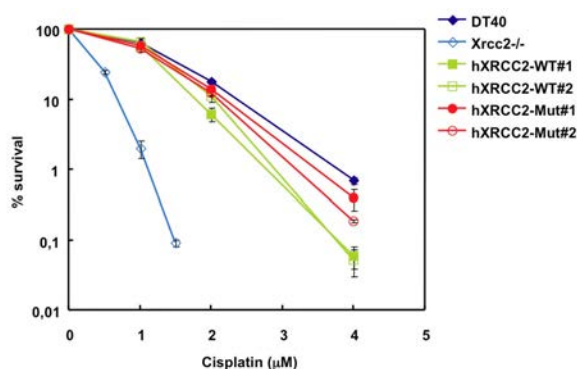


図 8 DT40 細胞株を用いた *XRCC* 遺伝子の R188H 変異のシスプラチン耐性に与える影響の評価

(2) 研究成果の今後期待される効果

DNA 修復遺伝子群の SNP は、放射線感受性、薬剤感受性に大きな影響を与える。今回

XRCC2 遺伝子とシスプラチンの組み合わせで示されたように、この R188H を持つ個人は、シスプラチンによる化学療法に耐性を示す可能性があり、より高濃度での薬剤使用が可能ではないかと考えられる。本研究で同定された DNA 修復遺伝子の SNP のうちアミノ酸変異を伴うものを、DT40 細胞株を用いたこのアッセイ系で網羅的に機能評価できれば、抗がん剤への耐性に関わる SNP を迅速で確実に探索することが可能である。こういった情報をデータベース化することによって、がん患者への投薬量の決定に客観的な指標を与えるものとして、将来的にテーラーメイド医療に大きく貢献することが期待される。

このサブテーマは、機能評価にどの程度の時間を要するか、この系がどの程度有効かを調べるためのパイロット実験を行ってきた。その結果上述の通り DT40 を用いたアッセイ系で *XRCC2* 遺伝子のアミノ酸変異を伴う SNP による遺伝子の活性の変化を検出することができた。今後ジェノタイピング結果から有意と考えられる SNP が検出された場合の機能評価に、この系が利用出来ることが明らかになったので、必要になれば再開することとして、2006 年度で終了した。

3.3 サブテーマ 3 高精度ハイスループット SNP 解析技術による SNP 部位の確認 (京都大学 辻本グループ)

(1) 研究実施内容及び成果

研究の目的

プライマー伸長反応を基盤とする高精度ハイスループット SNP 解析技術を用い、シーケンシングにより決定された SNP タイピングを確認するとともに、絞られた候補 SNP 部位の臨床応用を図る。現在までに辻本等は各種解析法を実際比較検討し、精度 (ほぼ 100%)、スループットにおいて秀でていいるのは現時点ではプライマー伸長反応を用いた方法論である事を確認している。本研究では、このプライマー伸長反応に基づくプラットフォームで、multiplex PCR、酵素増感法による蛍光検出の高感度化により現状のコストを更に大きく下げ、また 384well を基盤とするロボティクスによる自動化、ハイスループット化を図り、処理検体数は可変的且つ迅速に行える方法論を用いてシーケンシングにより決定された SNP タイピングを確認する。更に、この方法論の費用効果比が高く、かつハイスループットにできることが検証できた場合、絞られた候補 SNP 部位に関する多検体処理を行い、臨床応用を図る。

研究実施方法

以下のようなステップで研究を実施した。

1. 当研究室で共同開発されたプライマー設計用のソフトウェアを、multiplex で反応を行える PCR 断片を増やせるように改良し、それを用いてオリゴヌクレオチドをデザインする。
2. multiplex PCR 反応を行い、ソフトウェアの有効性、データの信頼性の評価、確認を、少数のマーカーと既にジェノタイプが決定されている健常者 DNA 検体を用いて試みる。
3. ジェノタイピングのコストダウンに向けて、酵素増感法による蛍光検出の高感度化を試みる。
4. この方法の有効性が立証されれば、造血器腫瘍を中心にジェノタイピングを行う。

現在の進捗状況と研究結果

SNP 領域の Multiplex PCR 増幅：至適プライマーのデザインを行うプログラムの構築はほぼ終了した。既に確立されている国際標準 DNA チップに関して、上記要素技術を用いて更に改良し微量臨床検体の解析のための更なる高感度化、高精度化、低コスト化の検討を行

った。その結果

1. SNP 領域の Multiplex PCR 増幅：至適プライマーのデザインを行うプログラムの改善を行い、約70%の成功率で10の multiplex が可能となった。
2. 高感度 DNA チップ開発を企業と共同で開始し、新たなマイクロアレイ基盤の開発に成功した。新型チップは従来の DNA チップと比べて約100倍の高感度であることが確認された。

(2) 研究成果の今後期待される効果

SNP ジェノタイピングは、ここ数年で試薬にかかる費用も劇的に下がり、GWAS を用いた大規模遺伝解析が世界各所で行なわれている。また、非常に近い将来それが全ゲノム塩基配列決定に取って代わられるであろうことも想像に難くない。しかしながら、それらの成果をふまえて、臨床レベルでのテーラーメイド医療が実施されるようになると、決まった数の SNP を如何に早く正確にタイピング出来るかが極めて重要となってくる。本研究で開発した、プライマー伸長法を用いるタイピング技術は、上のような現状で研究レベルでの大規模遺伝解析には利用されにくいと思われるが、病院ベースでの分子診断に大きな貢献をすると期待している。

このサブテーマは、マルチプレックス化のためのコンピュータプログラムの改良、より高い検出感度のチップの作成等、大規模タイピングを行える技術基盤の開発がほぼ終了したが、まだ劇的に安いコストで迅速、正確にタイピングが行えるだけのプラットフォームの整備が終了していない。またコストダウンに関しては、世界の企業から新たなテクノロジーが日夜開発、商品化されており、それらと競合して、比肩するものを開発するのは極めて困難であると思われる。そこで、2006年度で終了した。

§4 研究参加者

①松田グループ(人種間ゲノム多型情報の解析システムの構築の研究)

	氏名	所属	役職	研究項目	参加時期
○	松田 文彦	京都大学大学院医学研究科	教授	研究の総括及びデータベースの構築	H15.10～H21.3
	山田 亮	東京大学医学研究所	准教授	データの統計解析	H18.4～H21.3
	千田 昇平	京都大学大学院医学研究科	特任助手	ジェノタイピングチーム責任者	H18.4～H19.3
	角谷 寛	京都大学大学院医学研究科	准教授	臨床情報管理責任者	H19.4～H21.3
	高橋 めい子	京都大学大学院医学研究科	助教	ジェノタイピングチーム責任者	H19.4～H21.3
*	Renault, Victor	京都大学大学院医学研究科	CREST 研究員	データベースのプログラミング	H16.9～H17.11
	Renault, Victor	京都大学大学院医学研究科	産学官連携研究員	データベースのプログラミング	H19.4～H20.3
*	山口 昌雄	京都大学大学院医学研究科	CREST 研究補助員	データベースのプログラミング	H16.1～H19.9
*	豊田 幸子	京都大学大学院医学研究科	CREST 研究補助員	SNP 同定及びジェノタイピング	H16.4～H17.3
	平谷 仁美	京都大学大学院医学研究科	技術補佐員	SNP 同定及びジェノタイピング	H17.4～H18.4
*	吉住 明子	京都大学大学院医学研究科	CREST 研究補助員	コンピュータネットワークの管理と維持	H16.4～H18.3
*	平野 和子	京都大学大学院医学研究科	CREST 研究補助員	研究資料の整理及びデータ入力	H19.6～H20.3
	小久保 美紀	京都大学大学院医学研究科	教務補佐員	SNP 同定及びジェノタイピング	H17.11～H21.3
	Yovo-Vasilescu, Justine	京都大学大学院医学研究科	教務補佐員	SNP 同定及びジェノタイピング	H17.11～H19.6
	Palomares, Marie-ange	京都大学大学院医学研究科	教務補佐員	SNP 同定及びジェノタイピング	H19.4～H21.3
	廣澤 桂	京都大学大学院医学研究科	大学院生		H18.4～H20.3
	大東 賢英	京都大学大学院医学研究科	大学院生		H18.4～H20.3
*	Ratanajaraya, Chanavee	京都大学大学院医学研究科	大学院生 / CREST 研究補助員	SNP 同定及びジェノタイピング	H20.4～H21.3
*	景山 由貴	京都大学大学院医学研究科	CREST チーム事務員		H16.4～H19.3
*	金澤 雅美	京都大学大学院医学研究科	CREST チーム事務員		H19.4～H21.3

②園田グループ(多型・変異をもつヒト DNA 修復遺伝子機能の迅速同定法の開発の研究)

	氏名	所属	役職	研究項目	参加時期
○	園田 英一朗	京都大学大学院医学研究科	准教授	DNA 修復遺伝子群の機能解析	H15.10～H19.3

	武田 俊一	京都大学大学院医学研究科	教授	DNA 修復遺伝子群の機能解析	H18.4～H19.3
*	Danoy, Patrick	京都大学大学院医学研究科	CREST 研究員	遺伝子の細胞への導入実験	H15.10～H16.6

③辻本グループ(高精度ハイスループット SNP 解析技術による SNP 部位の確認の研究)

	氏名	所属	役職	研究項目	参加時期
○	辻本 豪三	京都大学大学院薬学研究科	教授	高効率 SNP システムの検証	H15.10～H19.3
	松本 治	京都大学大学院薬学研究科	准教授	プライマー伸長反応 probe のデザイン	H15.10～H19.3
	三輪 嘉尚	京都大学大学院薬学研究科	講師	プライマー伸長反応による SNP 検証	H15.10～H19.3

④白川グループ(多型・変異をもつヒト DNA 修復遺伝子機能の迅速同定法の開発の研究)

	氏名	所属	役職	研究項目	参加時期
○	白川 太郎	京都大学大学院医学研究科	教授	アレルギー免疫疾患関連遺伝子の同定	H15.10～H18.3
*	毛 暁全	京都大学大学院医学研究科	CREST 研究員	サンプルの収集	H16.4～H17.4

⑤Lathrop グループ(多因子型遺伝病の多人種間の比較解析の研究)

	氏名	所属	役職	研究項目	参加時期
○	Lathrop, Mark	Centre d'Etude du Polymorphisme Humain	Scientific Director	ジェノタイプ結果の統計解析	H15.10～H21.3
	Matsuda, Fumihiko	Centre National de Genotypage	Head of Gene	遺伝解析結果の統計解析	H15.10～H20.3
	Zelenika, Diana	Centre National de Genotypage	Project Manager	SNP 同定結果の解析	H16.3～H17.6
	Heath, Simon	Centre National de Genotypage	Head of Statistics	ジェノタイプマーカーの選択と結果の解析	H15.10～H21.3
	Boone, Claire	Centre National de Genotypage	Team Leader	実験チームのマネージメント	H17.11～H19.3
	Derbois, Celine	Centre National de Genotypage	Assistant Engineer	SNP 同定結果の解析	H15.10～H20.3
	Dulary, Cecile	Centre National de Genotypage	Senior Technician	DNA 修復遺伝子の SNP 同定とジェノタイピング	H15.10～H20.3
	Meziani, Roubila	Centre National de Genotypage	Senior Technician	免疫関連遺伝子の SNP 同定とジェノタイピング	H16.3～H21.3
	van den Broeck, Dominique	Centre National de Genotypage	Senior Technician	免疫関連遺伝子の SNP 同定とジェノタイピング	H20.3～H21.3

§5 招聘した研究者等

なし

§6 成果発表等

(1)原著論文発表 (国内(和文)誌 0 件、国際 (欧文) 誌 18 件)

1. Do, H., Vasilescu, A., Diop, G., Hirtzig, T., Coulonges, C., Labib, T., Heath, S., Spadoni, J. L., Therwath, A., Lathrop, M., Matsuda, F. and Zagury, J. -F. (2006) Associations of the IL2Ralpha, IL4Ralpha, IL10Ralpha, and IFN (gamma) R1 cytokine receptor genes with AIDS progression in a French AIDS cohort. *Immunogenetics*. **58**, 89-98.
2. Plancoulaine, S., Gessain, A., Tortevoeye, P., Boland-Auge, A., Vasilescu, A., Matsuda, F. and Abel, L. (2006) A major susceptibility locus for HTLV-1 infection in childhood maps to chromosome 6q27. *Hum. Mol. Genet.* **15**, 3306-3312.
3. Ami, Y., Shimazui, T., Akaza, H., Uematsu, N., Yano, Y., Tsujimoto, G., Uchida, K. (2005) Gene expression profiles correlate with the morphology and metastasis characteristics of renal cell carcinoma cells. *Oncol. Rep.* **13**: 75-80
4. Nagino, K., Nomura, O., Takii, Y., Myomoto, A., Ichikawa, M., Nakamura, F., Higasa M., Akiyama, H., Nobumasa, H., Shiojima, S., Tsujimoto, G. (2006) Ultrasensitive DNA chip: Gene expression profile analysis without RNA amplification. *J. Biochem. (Tokyo)* **139**: 697- 703
5. Yoshimura, A., Nishino, K., Takezawa, J., Tada, S., Kobayashi, T., Sonoda, E., Kawamoto, T., Takeda, S., Ishii, Y., Yamada, K., Enomoto, T., and Seki, M. (2006) A novel Rad18 function involved in protection of the vertebrate genome after exposure to camptothecin. *DNA Repair (Amst)* **5**, 1307-1316. Epub 2006 Aug 22.
6. Kawanishi, H., Takahashi, T., Ito, M., Watanabe, J., Higashi, S., Kamoto, T., Habuchi, T., Kadowaki, T., Tsujimoto, G., Nishiyama, H., Ogawa, O. (2006) High throughput comparative genomic hybridization array analysis of multifocal urothelial cancers. *Cancer Sci.* **97**: 746- 752
7. Danoy, P., Sonoda, E., Lathrop, M., Takeda, S. and Matsuda, F. (2007) A naturally occurring genetic variant of human XRCC2 (R188H) confers increased resistance to cisplatin-induced DNA damage. *Biochem. Biophys. Res. Commun.* **352**, 763-768.
8. Vasilescu, A., Terashima, Y., Enomoto, M., Heath, S., Poonpiriya, V., Gatanaga, H., Do, H., Diop, G., Hirtzig, T., Auewarakul, P., Lauhakirti, D., Sura, T., Charneau, P., Marullo, S., Therwath, A., Oka, S., Kanegasaki, S., Lathrop, M., Matsushima, K., Zagury, J. -F. and Matsuda, F. (2007) A haplotype of the human CXCR1 gene protective against rapid disease progression in HIV-1+ patients. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3354-3359.
9. Yamada, R. and Matsuda, F. (2007) A novel method to express SNP-based genetic heterogeneity, Psi, and its use to measure linkage disequilibrium for multiple SNPs, D(g), and to estimate absolute maximum of haplotype frequency. *Genet. Epidemiol.* **31**, 709-726. Epub 2007 May 16.
10. Murakawa, Y., Sonoda, E., Barber, L.J., Zeng, W., Yokomori, K., Kimura, H., Niimi, A., Lehmann, A., Zhao, G.Y., Hochegger, H., Boulton, S.J., and Takeda, S. (2007) Inhibitors of the proteasome suppress homologous DNA recombination in mammalian cells. *Cancer Res.* Sep **67**, 8536-8543
11. Thein, S. L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., Silver, N., Gerovasilli, A., Ping, C., Yamaguchi, M., Wahlberg, K., Ulug, P., Spector, T. D., Garner, C., Matsuda, F., Farrall, M. and Lathrop, M. (2007) Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11346-11351.
12. Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H., Best, S., Spector, T. D., Farrall, M., Lathrop, M. and Thein, S. L. (2007) A

QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nature Genet.* **39**, 1197-1199. Epub 2007 Sep 2.

13. Ridpath, J.R., Nakamura, A., Tano K., Luke A.M., Sonoda, E., Arakawa, H., Buerstedde, J.M., Gillespie, D.A., Sale, J.E., Yamazoe, M., Bishop, D.K., Takata, M., Takeda, S., Watanabe, M., Swenberg, J.A., and Nakamura, J. (2007) Cells deficient in the FANC/BRCA pathway are hypersensitive to plasma levels of formaldehyde. *Cancer Res.* **67**, 11117-11122
14. Michiels, S., Danoy, P., Dessen, P., Bera, A., Boulet, T., Bouchardy, C., Lathrop, M., Sarasin, A. and Benhamou, S. (2007) Polymorphism discovery in 62 DNA repair genes and haplotype-associations with risks for lung, and head and neck cancers. *Carcinogenesis.* **28**, 1731-1239
15. Tocharoentanaphol, C., Promso, S., Zelenika, D., Lowhnoo, T., Tongsimas, S., Sura, T., Chantratita, W., Matsuda, F., Mooney, S. and Sakuntabhai, A. (2008) Evaluation of resequencing on number of tag SNPs of 13 atherosclerosis-related genes in Thai population. *J. Hum. Genet.* **53**, 74-86.
16. Hung, R.J., McKay, J.D., Gaborieau, V., Boffetta, P., Hashibe, M., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., Lissowska, J., Rudnai, P., Fabianova, E., Mates, D., Bencko, V., Foretova, L., Janout, V., Chen, C., Goodman, G., Field, J.K., Liloglou, T., Xinarianos, G., Cassidy, A., McLaughlin, J., Liu, G., Narod, S., Krokkan, H.E., Skorpen, F., Elvestad, M.B., Hveem, K., Vatten, L., Linseisen, J., Clavel-Chapelon, F., Vineis P., Bueno-de-Mesquita, H.B., Lund, E., Martinez, C., Bingham, S., Rasmuson, T., Hainaut, P., Riboli, E., Ahrens, W., Benhamou, S., Lagiou, P., Trichopoulos, D., Holcátová, I., Merletti, F., Kjaerheim, K., Agudo, A., Macfarlane, G., Talamini, R., Simonato, L., Lowry, R., Conway, D.I., Znaor, A., Healy, C., Zelenika, D., Boland, A., Delepine, M., Foglio, M., Lechner, D., Matsuda, F., Blanche, H., Gut, I., Heath, S., Lathrop, M. and Brennan, P. (2008) A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.* **452**, 633-637.
17. Danoy, P., Michiels, S., Dessen, P., Pignat, C., Boulet, T., Monet, M., Bouchardy, C., Lathrop, M., Sarasin, A. and Benhamou, S. (2008) Variants in DNA double-strand break repair and DNA damage-response genes and susceptibility to lung and head and neck cancers. *Int J Cancer.* **123**, 457-463
18. McKay, J.D., Hung, R.J., Gaborieau, V., Boffetta, P., Chabrier, A., Byrnes, G., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., Lissowska, J., Rudnai, P., Fabianova, E., Mates, D., Bencko, V., Foretova, L., Janout, V., McLaughlin, J., Shepherd, F., Montpetit, A., Narod, S., Krokkan, H.E., Skorpen, F., Elvestad, M.B., Vatten, L., Njølstad, I., Axelsson, T., Chen, C., Goodman, G., Barnett, M., Loomis, M.M., Lubiński, J., Matyjasik, J., Lener, M., Oszutowska, D., Field, J., Liloglou, T., Xinarianos, G., Cassidy, A.; EPIC Study, Vineis, P., Clavel-Chapelon, F., Palli, D., Tumino, R., Krogh, V., Panico, S., González, C.A., Ramón Quirós, J., Martínez, C., Navarro, C., Ardanaz, E., Larrañaga, N., Kham, K.T., Key, T., Bueno-de-Mesquita, H.B., Peeters, P.H., Trichopoulou, A., Linseisen, J., Boeing, H., Hallmans, G., Overvad, K., Tjønneland, A., Kumle, M., Riboli, E., Zelenika, D., Boland, A., Delepine, M., Foglio, M., Lechner, D., Matsuda, F., Blanche, H., Gut, I., Heath, S., Lathrop, M. and Brennan, P. (2008) Lung cancer susceptibility locus at 5p15.33. *Nat Genet.* 2008 Nov 2. [Epub ahead of print]

(2)学会発表(国際学会発表及び主要な国内学会発表)

① 招待講演 (国内会議 4 件、国際会議 4 件)

1. Fumihiko Matsuda: “Explorer la diversité génomique chez l’homme” 8ème Carrefour Européen des Biotechnologies. Marseille, France 2004. 10. 28-29.
2. 松田文彦: 「免疫系の疾患とゲノム医学」分子予防環境医学研究会シンポジウム「ゲノム医科学の予防医学への統合」2004年12月20日

3. 松田文彦：「がんの素因と放射線感受性に関するゲノム疫学」 放射線疫学情報シンポジウム 2005年3月9日-10日
4. Fumihiko Mstsuda: “A trans-ethnic genomics study of DNA repair genes” The Second Nagasaki Symposium of International Consortium for Medical Care of Hibakusha and Radiation Life Science. Nagasaki University, Nagasaki 26-27 July, 2006
5. Fumihiko Matsuda: “Genomic study of human diseases” France/Japan Conference of Collaborative Opportunities in Biomedical Research. La Fondation del Duca - Institut de France, Paris France. 8-9 February, 2007
6. 松田文彦：「ヒト多重遺伝子疾患の多人種間比較による遺伝因子同定」第27回日本医学学会総会 シンポジウム「ヒトゲノムと疾患」2007年4月6日-8日
7. 松田文彦： “Trans-ethnic genetic study of human multigenetic disorders” 日本循環器学会 福岡県福岡市 2008年3月29日
8. Fumihiko Matsuda: “Genomic study of human multigenetic disorders” The 11th Kyoto University International Symposium 2008 “Fronier Bioscience in Modern Medicine”. Fudan University, Shanghai, China. 9-11 October, 2008

② 口頭発表 (国内会議 2 件、国際会議 1 件)

1. Hiroaki Yabuuchi, Yasushi Okuno, Susumu Katsuma, Gozoh Tsujimoto: “Sense-Antisense Transcriptome Network: Hidden relationships among the transcripts” HUGO's 10th Human Genome Meeting 国立京都国際会館、2005年4月18日-21日
2. Tsujimoto G. “GPR120 and GPCR Family for Free Fatty Acids” The 2006 Keystone Symposium on G Protein-Coupled Receptors: Evolving Concepts and New Techniques. 12-17 February, 2006.
3. Vasilescu, A., Saito, M., Yasunaga, J., Matsuoka, M., Yamada, R. and Matsuda, F. “Genomic study of HTLV-I infection in Japanese population” 日本人類遺伝学会 東京都 2007年9月13日

③ ポスター発表 (国内会議 3 件、国際会議 0 件)

1. 奥野恭史、種石慶、土屋創健、王質輝、辻本豪三「ゲノム創薬支援システムの開発」日本薬学会第126年会 仙台市、2005年3月28日-30日
2. 廣澤桂、山田亮、松田文彦「一塩基多型(SNP)を用いた有意検定における、観測ジェノタイプ情報の組み込みに関する検討」日本人類遺伝学会第52回 東京都 2007年9月13日
3. 大東賢英、山田亮、松田文彦「大量の SNP をマーカーとして用いた日本人関節リウマチのケースコントロール解析」日本人類遺伝学会第52回 東京都 2007年9月13日

(3)特許出願

①国内出願 (1 件)

1. 発明の名称：データ処理装置、データ処理プログラム、それを格納したコンピュータ読み取り可能な記録媒体、およびデータ
発明者：辻本 豪三、奥野 恭史、梁 智允、種石 慶
出願人：国立大学京都大学

出願日：2005年6月30日
出願番号：2005-192675

②海外出願 (0件)

(4)受賞等

①受賞

なし

②新聞報道

2006年4月3日 産経新聞「人種を超えて遺伝子解析」

2006年8月17日 京都新聞「生命科学は今」 遺伝子の個人差

③その他

2007年1月1日 フランス国立医学健康科学研究機構(INSERM)海外研究ユニット U.852 が、研究代表者の所属するゲノム医学センターに設置された。ユニットの設置期間は8年間で、総額80万ユーロの研究予算で「ヒト複合遺伝性疾患の比較遺伝解析」を行なう。

(5)その他特記事項

なし

§7 研究期間中の主な活動

なし

§8 結び

ゲノム科学は恐ろしい勢いで進歩しており、時にはその中にどっぷり浸かっている人間でさえ、そのスピードについていくのが困難である。プロジェクトを始めた2003年と現在を比較すると、おそらくジェノタイピングのスループットは3桁以上上がり、また当時は少数の候補遺伝子を選んで細々と行っていた複合遺伝性疾患の感受性遺伝子検索は、数十万個のマーカーを搭載したSNPアレイによるゲノムワイド関連解析(GWAS)にほぼとって代わられた。そういった時代の急速な変化のために、開始当初考えていた候補遺伝子解析はあっという間に前時代の遺物となり、現在世界の一線でのぎを削っているグループが標準的な戦略として採用している何千人もの検体を用いた大規模なGWASも、ここ数年の間に全ゲノムシーケンスに置き換わるものと予想される。

こういった急速なテクノロジーの変化を過小評価していた自分の見通しの甘さがゆえに、プロジェクトの中間評価で、方針変更の助言をいただけなかったら、惨憺たる結果となっていたと思うと、今でも背筋が寒くなる。また逆にあのタイミングでたいへん貴重な助言をいただいた、笹月総括をはじめ領域アドバイザーの先生方には、どれだけ感謝してもきれない思いである。ことに笹月先生には、遅々として進まないプロジェクトにさぞかし気をもまれ、たいへんなご心配をおかけしてしまったにもかかわらず、いつも適切な御助言とともに暖かく見守っていただき、心よりお礼申し上げます。

幸いプロジェクト後半に開始した関節リウマチと肺がんのGWASが、たいへんいい結果に結びつきつつあり、2009年3月でプロジェクトが終了してしまうのが、たいへん残念である。しかしながら、近い将来必ず大きな実りが得られることを信じて、今後も全力で研究を続け、先生方のご恩に報いたいと思っている。

最後に、科学技術振興機構の皆様、特にテラーメイド事務所のスタッフの皆様方に、深くお礼申し上げます。