

定量的一塩基多型解析技術の開発と医療への応用

◆田平 知子¹⁾ ◆久木田 洋児²⁾ ◆堀内 孝彦³⁾ ◆林 健志¹⁾

1) 九州大学生体防御医学研究所 2) 大阪府立成人病センター研究所 3) 九州大学病院

<研究の目的と進め方>

ポストシーケンシング時代のゲノム科学研究では、多因子性遺伝性疾患の関連解析による原因遺伝子探索が最重要課題である。この探索には、ゲノムワイドなハプロタイプ構造情報に基づいた膨大な数のSNPについて、極めて多数の被検試料を用いた頻度解析が要求される。このためには、従来の各個的、散発的な研究とは異なり、システムティックな実験設定とデータ管理が要求される。本研究の目的は、(1)既に我々が確立した自動化された一塩基多型(SNP)アレル頻度高効率高精度決定法である「pooled DNAのPLACE-SSCP解析法」を発展させ、多数の多因子性遺伝性疾患の原因遺伝子を大規模に探索することを現実的なものとする方法論を確立するとともに、(2)関連解析に必要な日本人ゲノムの多様性、地域差を解明し、(3)ハプロタイプ多様性に関する情報基盤を確立する。また、(4)これらの方法論及び情報基盤を利用して、自己免疫疾患をはじめとする多因子性遺伝性疾患の候補遺伝子の大規模な関連解析を行い、疾患の要因となる遺伝的背景を解明する。これによって、多因子性遺伝性疾患の適切な診断、新治療法の開発、及び予防に寄与することを目標とする。

<研究開始時の研究計画>

これまで我々はPCR産物をポストラベルシ、蛍光キャピラリーアレイ自動シーケンサーを用いてSSCP解析する手法(PLACE-SSCP法)を確立し、さらにDNAプールを対象にしたPLACE-SSCP解析によるSNPアレル頻度算出法を開発してきた。また、この方法を利用して遺伝子転写開始点領域のSNP約10,000個を網羅的に解析しdbQSNPデータベースとして公開してきた。この成果は、「pooled DNAのPLACE-SSCP解析(定量的SSCP解析)」によるSNPアレル頻度決定法が極めて信頼度が高く、ほとんどのSNPについてこの方法が適用可能であることを証明するものである。従って患者群・対照群でのSNPアレル頻度を比較する関連解析にこの方法は有用であると考えられる。これまでの研究で、高度にシステム化された実験・データ管理システム「dbQSNP」を開発した。これはUNIXサーバー上で稼動するリレーショナルデータベースを用いた大規模なシステムである。これを関連解析の手法として一般に普及させるのは不適當であるので、広く疾患の関連解析にこのシステムを適用して、多数の多因子疾患の原因遺伝子探索を効率よく行うために、Windows上で稼動する上記システムの軽量版の開発を行う。一方、日本人でのSNPアレル頻度の多様性を検討し、関連解析における対照群選択のための情報基盤を整える。またゲノムワイド関連解析の効率化に必要な情報基盤として、日本人ハプロタイプ構造に関する情報を整備する。さらに先行研究として、典型的な多因子性自己免疫疾患である全身性エリテマトーデス(SLE)等の関連解析を多数の候補遺伝子領域にあるSNPを用いて行い、疾患原因遺伝子を追究する。

<研究期間の成果>

1. DNAプールの定量的SSCP解析によるアレル頻度計測のためのソフトウェア開発

DNAプールの定量的SSCP解析を広く関連解析研究に応用するために、「dbQSNP」の機能を取り入れてWindows上で独立に作動するSNP同定、アレル頻度定量解析ソフトウェア「QSNP lite」を開発した。これは、SSCP解析によるSNPタイピング及びアレル頻度決定と、シーケンシングによるSNP配列の同定・確認を統合して行うものである。また、96ウェルプレート単位で多数サンプルの解析を行うことを可能とするための実験手順の設定、キャピラリーアレイシーケンサーからのデータ管理等を行う機能を有している。このソフトウェアは、SSCP解析における波形解析にはこれまでの研究で独自開発した「QUISCA」モジュール(Higasa et al. 2002)を用いるが、さらに各アレルのピークが重なっている場合に幾何学的に分離してアレル頻度を算出する機能も有する。シーケンシング解析はWindows内にCoLinuxの仮想ディスクを組み込み、その上でベースコール・整列のための「PhredPhrap」(<http://www.phrap.org/phredphrapconsed.html>)及び配列変化検出のための「PolyPhred」(<http://droog.gs.washington.edu/polyphred/>)を実行することによる。また挿入・欠失変異を高い確度で判定する独自開発の機能も組み込んでいる。

このソフトウェアを利用して多数のSNPについてプールDNAのPLACE-SSCP解析によりアレル頻度を推定し、そのプールを構成する各検体のタイピングから得られたアレル頻度と比較することにより推定の精度を検討し、両者が極めて高い相関を示すことを確認した(Tahira et al., 2006)。またこのシステムを用いてプールDNAの定量的SSCP解析によりSNPアレル頻度を決定するプロトコルを総説として発表した(Tahira et al. 2009)。

2. 定量的SNP解析技術を利用した病因遺伝子解析

1) DNAプールの定量的SSCP解析を利用した候補遺伝子解析

自己免疫疾患である全身性エリテマトーデスの発症機構は未だ不明で、環境要因とともに遺伝的要因の関与が考えられる。この遺伝的要因を解明するために関連解析を行った。先行研究として「dbQSNP」システムを用いて免疫系のシグナル伝達やアポトーシスに関連する53候補遺伝子の全エクソンおよびプロモーター領域のSNPを患者検体から探索し、患者群および対照群のDNAプールのアレル頻度を定量し関連解析を行った。これにより補体C3をコードする遺伝子内に疾患との関連の強いSNPを検出した。この関連を個別サンプルのシーケンスによるタイピングによっても確認し、このSNPのリスクアレルの数と血中C3濃度の間の相関を検出し、このSNPを含むハプロタイプがC3の発現あるいは安定性にかかわっている可能性を示した(Miyagawa et al. 2008)。

次に欧米での先行研究でSLEとの関連が示唆された遺伝子について同様にプールDNAを用いた関連解析を行い、転写因子

IRF5 遺伝子領域の SNP が疾患と強い関連を示すことを見出した。個別サンプルのタイピングによって、欧米人とアジア人ではリスクハプロタイプには違いがあることを明らかにした。また、IRF5 プロモーターの上流の転写因子 Sp1 結合部位の繰り返し数が 4 であるアレルをホモに持つ人は 3 をホモに持つ人に比べリスクが高くなるという報告と一致する結果を得た。さらに HapMap のアジア人サンプルでこの繰り返し配列をタイピングし、GENEVAR データ (<http://www.sanger.ac.uk/humgen/genevar/>) より得た各サンプルの IRF5 発現データと比較することにより、繰り返し数が多いハプロタイプの発現が高いことを見出し、IRF5 発現レベルが SLE 易罹患性と関連すると考察した。

2) 定量的SSCP解析を利用したがん組織での遺伝子異常の解析

「QSNPlite」ソフトウェアを用いた PLACE-SSCP 解析は SNP アレルの量比を正確に決定できるので、癌細胞における loss of heterozygosity (LOH) の検出にも利用できる。そこでこの解析法を脳腫瘍組織での LOH 解析に応用した。その結果、正常組織が 70% を占めるような腫瘍組織でも、LOH を的確に検出可能であることを明らかにし、グリオーマにおける高頻度 LOH 領域を決定した (Hata et al., 2006)。またこの方法をさらに発展させ多数の髄芽腫検体の LOH 解析により共通欠失部位を同定した (Guan et al., 2008)。このような LOH をゲノムワイドに解析するため、マイクロアレイを用いた定量的解析も行った。この場合は、正常組織の混入があると定量性が著しく損なわれるのでレーザーマイクロダイセクション法により切り出した腫瘍細胞から抽出したゲノム DNA をマイクロアレイで解析した。これによりコピー数に変化しないタイプの LOH、つまり片側のアレルが倍加しているものを高頻度に検出した。これらをマイクロサテライト解析および染色体 FISH 解析により確認し、このような特異な染色体変化が癌化に関わっていることを示した (Kuga et al. 2008)。

3) 遺伝性眼科疾患の変異解析

家族性滲出性硝子体網膜症 (FEVR) の多数の家系での原因遺伝子突然変異探索を行い、細胞外シグナル分子 *norrin* 及びこれの受容体—シグナル伝達系蛋白質である *frizzled 4* (Wnt シグナルの細胞表面受容体) をコードする *FDZ4*、共受容体をコードする *LRP5*、の 3 つの共役複合体の遺伝子にそれぞれ複数の突然変異・多型を同定した (Qin et al., 2005; Kondo et al., 2007)。変異の違いにより疾患重症度の相違が見られる機構を解明するため、アミノ酸レベルでの変異・多型が Wnt シグナル標的遺伝子の発現へ与える影響をレポーターアッセイにより定量した。その結果、Wnt シグナル古典的経路の阻害と表現型が必ずしも相関しないことを見出し発症要因が複合的であると結論した (Qin et al., 2008)。

3. プールDNAを用いたゲノムワイド関連解析 (P-GWAS)

近年の DNA マイクロアレイ技術の進歩により、多因子疾患の原因遺伝子を網羅的に探索するゲノムワイド関連解析は現実的なものとなってきた。しかし、既成のシステムを利用して非常に多数の検体を解析するのはコスト面での限界がある。大規模ゲノムワイド関連解析を低コストで行う方法論を確立するため、まず予備実験として、Affymetrix 社の 100K 及び 500K SNP アレイを用いてプール DNA から高精度に SNP アレル頻度を推定する実験方法及びデータ処理を検討した。このために上記アレイの全ての SNP に関するジェノタイプが Affymetrix 社によって決定・公開されている Caucasian 及び African American の DNA を Coriell 研究所から購入し、それぞれ 42 人からなる 2 種類の DNA プー

ル試料とした。これを 100K 及び 500K アレイを用いて多数回解析し、蛍光強度シグナルデータを得た。これらをもとにアレル間相対シグナル強度 (relative allele signal, RAS) の推定法、RAS からアレル頻度への変換の適否、その方法等を検討した。これにより Perfect match (PM) プローブのみを用いて計算した RAS の値を用いたほうが Miss Match (MM) プローブで補正した場合より実験の再現性がよいことを見出した。またアレル間の蛍光シグナルの差を補正しなくても関連解析では患者群と対照群の相対的な差がわかればよいので問題ないと考えた。

SLE 患者群および対照群の DNA プールを被検材料とし、SLE の P-GWAS を行った。まず、Affymetrix 500K アレイによる定量的 SNP 解析で 1 次スクリーニングを行った。測定的主要誤差はハイブリダイゼーション段階で生じるので、同一プールの複数回定量により誤差を軽減するために各 6 回解析した。両群の比較は RAS により行った。この段階で米国 TGEN のグループにより、500K アレイを用いた関連解析の手法が報告され、我々の予備実験と同様に PM プローブのみから算出した RAS を使用して患者群・対照群の比較を行うことが提唱された。その際にそれぞれの RAS のクラスターがどのぐらい分離しているかの指標であるシルエットスコアで SNP をランク付けする方法が最も有効であるという結果が得られていたので、我々もこの方法を採用した。このほかに各マーカのセンスプローブ、アンチセンスプローブのそれぞれの RAS の中央値を比較して z-スコアを算出し、それによりランク付けする方法も行った。マイクロアレイを用いた解析ではノイズによる偽陽性が大きな問題となる。これを除外するためにいずれの統計量についてもスライディングウィンドウ解析を行った。この解析では複数の SNP が連鎖不平衡にある場合にその領域のランクが上がるので検出力が向上する。連鎖不平衡の程度は領域により異なるので、それらを検出するために 3-15 マーカーまでのウィンドウサイズを設定した。これにより、患者群と健常者群のプールでアレル頻度が異なると予想される複数の領域が検出された。

この中からさらに偽陽性のシグナルを除外するために、より精度の高い定量が可能である PLACE-SSCP 法を 2 次スクリーニングとして用いて各領域でスコアの高かった SNP についてアレイ解析によるスクリーニングに用いたものと同じ DNA プールの SNP アレル頻度を測定し、疾患と関連が強い領域をさらに絞り込んだ。

これら 2 段階の検討で非常に強い関連が示された複数の SNP について TaqMan アッセイによる個別タイピングによって最終的に疾患との関連を確認した。最も強く疾患と関連していたのは IKZF1 (Ikaros) 遺伝子の上流で、このほかに PRDM1 (Blimp1) 遺伝子下流領域、HLA-G 下流領域、さらに欧米におけるゲノムワイド関連解析で SLE との関連が検出された TNFAIP3(A20) 遺伝子領域、BLK 遺伝子が上位に検出された。IKZF1 遺伝子の上流、PRDM1 遺伝子下流領域、HLA-G 下流領域、の 3 領域の SNP については東京大学の山本一彦博士との共同研究で別の地域のサンプルでも関連解析を行い、すべてが再現性よく疾患との関連を示すことを確認した。特に IKZF1 遺伝子の上流の SNP はメタアナリシスによっても強い関連を示した (odds ratio 1.5, $P=3 \times 10^{-13}$)。現在、周辺領域のマーカー数を増やし更に詳細な解析を行っている。

今回の解析では SLE との強い関連がすでに報告されている STAT4 遺伝子、IRF5 遺伝子の SNP は検出されなかった。これらの SNP (およびこれらと強い連鎖不平衡にある SNP) は使用したアレイに搭載されておらず、従って、今回検出されなかったのはマーカーの搭載がまだ不十分であるためである。またこのほ

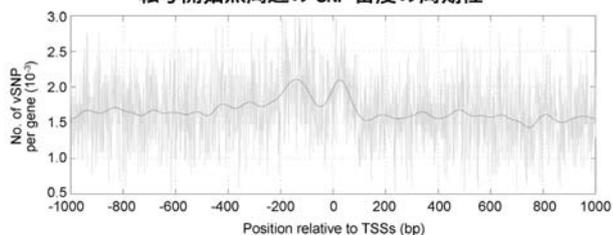
かに欧米の先行研究で強い関連が検出されているにもかかわらず今回の解析で検出できなかった多数の領域について定量的SSCP解析で関連を検討した結果、関連が弱いことが確認された。これらは偽陰性ではなく感受性遺伝子に人種差があることにより検出されなかったと考えられた。

以上の結果は、患者群および対照群のDNAプールをゲノムワイドアレイでのタイピングによる定量とPLACE-SSCP法による定量の2段階で関連解析する手法が疾患関連SNPを効率的に同定するのに有用であることを示した。また、SLEの疾患要因として、多数の遺伝子が関与していることを明らかにした。さらに、今回同定した感受性領域は、他の自己免疫疾患とも関連している。すなわち、IKZF1遺伝子の上流はクローン病との関連が、PRDM1遺伝子下流領域は関節リウマチとの関連が報告されている。このことから、未だその機能は不明であるが自己免疫疾患に共通のパスウェイに係わる遺伝子領域を検出していると考えられる。

4. 遺伝子転写開始点領域のSNPのアレル頻度と分布

これまでの研究で遺伝子転写開始点領域のSNPを同定し、また日本人のプールDNAと西欧人のプールDNAとを対象としてアレル頻度を定量的SSCP解析により推定し、dbQSNPデータベースとして公開している。ここで見出されたSNPについて転写開始点を起点とした場合の分布およびアレル頻度の人種間での違いを報告した(Tahira et al., 2005)。このとき、SNP密度に周期性がある傾向に気付いた。この周期性を確認しその原因を明らかにするために、更に米国のdbSNPデータベースのSNPを対象として詳細に解析したところ、この周期は146塩基であり、CpGアイランドを有する多くの遺伝子の転写開始点付近に特徴的に認められた。この結果より、生殖細胞系列での転写開始点近傍のヌクレオソーム構造が変異率あるいは変異の定着に関与している可能性を提示した(Higasa & Hayashi, 2006)。

転写開始点周辺のSNP密度の周期性



5. 日本人集団内でのSNPアレル頻度の地域差の検討

関連解析では対象とする集団の遺伝的均質性が常に問題となる。そこで、がんコホート研究で収集された、異なる7地域で収集された日本人集団のDNAを用いてSNPアレル頻度の違いがあるかどうかを調べた。即ち各集団それぞれ400人分のDNAプールを作製し、PLACE-SSCP法により定量した。細菌感染感受性に関わることが判明しているMBL遺伝子上のSNPに関しては感受性アレルの頻度に地域差がないという結果を得た(Horiuchi et al., 2005)。

6. 全胎状奇胎を用いた全ゲノムハプロタイプの直接決定

1) 確定的SNPハプロタイプの決定およびデータベース構築

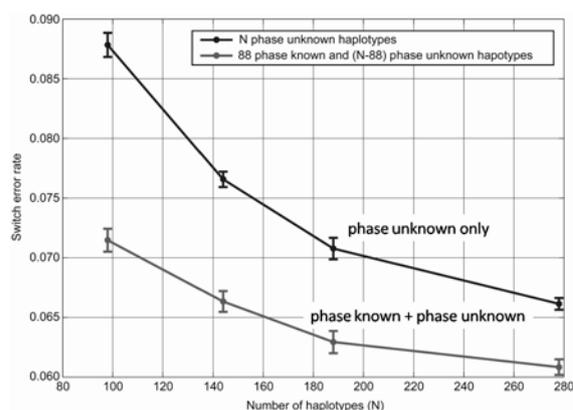
多因子疾患の原因遺伝子をゲノムワイド関連解析により同定するために必要な全ゲノムのハプロタイプを決定する目的で国際ハップマッププロジェクトが遂行された。しかし、このプロジェクトでは2倍体細胞に由来するDNAを用いて決定したSNPジェ

ノタイプから間接的にハプロタイプを推定している。特に、東アジア人(日本人及び中国人)の試料に関しては、不特定個人45人由来のジェノタイプデータを基にハプロタイプを遺伝統計学的モデルのみに基づいて推定しているため、その結果は誤りを多く含む可能性がある。そこで我々は、単一精子由来のハプロイドゲノムを持つ全胎状奇胎を解析することにより、日本人ゲノムのハプロタイプを直接決定し、より精細な疾患の関連解析に耐えうるゲノム情報基盤を確立することを目指した。

まず74個の胎状奇胎を試料としてゲノムワイドに28万個のSNPについてジェノタイプングを行い、これにより得られた74セットのゲノムワイドな確定ハプロタイプ(D1 phase)をもとに情報学的解析を行い、日本人ゲノムのハプロタイプブロック構造を決定した。また、関連解析を効率よく行うためにタグSNPの選定を行った。これらの結果を高度検索機能を有するデータベース「D-Haplo-DB」(<http://orca.gen.kyushu-u.ac.jp>)に集約し、公開した(Kukita et al., 2005; Higasa et al., 2007)。種々のシミュレーションから、ここで得られたタグSNPを用いたゲノムワイド関連解析を行った場合、約60%の確率で原因遺伝子を相関係数(r^2)0.8以上として検出しようと推定した。

次に胎状奇胎100個をAffymetrix 500K SNPアレイでタイピングし、連鎖不平衡構造及びタグSNP情報をより精細にした(D2 phaseとして公開、計581K SNPs)。D2 phaseのタグSNPにより、頻度5%以上の未知SNPの80%を $r^2 > 0.8$ としてキャプチャーできた。また、連鎖不平衡構造がD1, D2間で変わらない領域も多く、それらの領域では既に連鎖不平衡解析に十分なSNP密度が達成されたと考えた。さらに、Affymetrix SNPアレイ6.0によりコピー数多型を含む約180万個のマーカーについてタイピングした。このうち約90万個のSNPタイピングの結果をD-Haploデータベース(D3 phase, 1M SNPs)に公開した。またディブロイドゲノムからハプロタイプを決定する際に、この確定的ハプロタイプの情報を用いることで、スイッチエラーが減少し推定精度が向上することを明らかにした(Higasa et al., 2009)。

確定ハプロタイプを含めたハプロタイプ決定によるスイッチエラーの減少



Illumina 1Mアレイによりコピー数多型(CNV)プローブを含む約107万個のマーカーについてもタイピングし、Affymetrix SNP 6.0アレイの結果とあわせて140万個のSNPからなるハプロタイプを決定した。

2) 確定的ハプロタイプの解析による自然選択の検出

全胎状奇胎DNAによる日本人確定ハプロタイプ構造(D-HaploDB)と国際HapMap計画による日本人推定ハプロタイプ構造とを比較した結果、大部分の領域では類似の構造であるが、HLA領域をはじめ、多くの領域で推定ハプロタイプの誤り(ス

イッチエラー)の頻度が高くなっており、またこのような領域の多くがこれまでの遺伝学的解析で自然選択を受けた領域として報告されたものと一致することを見出した(Higasa et al., 2009).

3) コピー数多型(Copy number variation: CNV)を含むハプロタイプの決定

CNVは近年疾患との関連が注目されているものであるが、ディプロイドゲノムでこれを検出するのは技術的に困難であり、またCNVのハプロタイプを決定するのも難しい。全胎状奇胎は、倍加したハプロイドゲノムをもつので、これを解析することによりCNVが感度良く検出されると考えられる。実際Illumina 1Mアレイと、Affymetrix SNP 6.0アレイの両方でCHMサンプルと通常のディプロイドサンプルを解析した結果を比較し、CHMサンプルではCNV領域がはるかに検出されやすいことを確認した。またディプロイドゲノムの場合のように2つのアレルのCNVセグメントの重なりにより範囲が不確定になることもない。そこで、CNVプローブを持つAffymetrix SNPアレイ6.0によりタイピングした結果のコピー数解析を行い、これにより日本人ゲノムのCNV領域(すべてのCNVセグメントを統合したもの)を決定した。このときクオリティコントロールを厳密に行い最終的に85個のCHM中に約7000個のCNVセグメント、1300個のCNV領域を同定した。欠失のほうが増幅より数多く、サイズは小さかった。また多くのCNV領域がセグメントゲノム重複(segmental duplication)の領域にあり、その場合はサイズが大きい傾向があった。McCarrollら(2008)によりHapMap計画の日本人サンプル(JPT)を同じアレイで解析して同定されたCNP(Copy number polymorphism, CNVのうち頻度が1%以上のものと定義)が報告されたので、その結果(頻度2%以上に限定)と比較したところ約4割は同じ領域に検出された。共通するCNVのサイズは両方のサンプルでよく相関しており、同じCNVを検出したことが示された。我々が見出したCNV領域の約6割は新しく見つかった領域であり、そのほとんどは頻度が低いものであった。

各ハプロタイプのCNVセグメントを相互の重なり具合からCNVイベント(CNVE)にまとめ、その周辺のSNPハプロタイプ構造を比較した。その結果、同じCNV領域のCNVEの間ではハプロタイプが有意に類似していた。この結果から、特定のハプロタイプでCNVEがおこりやすいことが判明した。この結果はSNPがほぼランダムに生じると対照的であり、CNVの形成機構の理解につながるものである。

<国内外での成果の位置づけ>

全胎状奇胎を用いて確定ハプロタイプを決定するのは世界で初めての試みであり、それをデータベース化して公開している意義は大きい。国内では日本人のゲノムワイドハプロタイプ構造に関する研究はほとんど行われておらず我々の研究は先行している。このため、以下の講演・講習を要請された。

- ・ゲノムリテラシー講座(科学技術振興機構 パイオインフォマティクス推進センター主催)2006年8月25日、東京
「決定的ハプロタイプの概要とそのデータベースD-HaploDBの利用法」林 健志、宮武克行、日笠幸一郎
- ・第5回国際バイオデータ相互運用性会議国際シンポジウム“遺伝多型解析研究の最先端”2007年9月26日、東京
「D-HaploDB: A database of genome-wide definitive haplotypes determined using complete hydatidiform mole」林 健志
- ・第3回バイオインフォマティクス研究者と医学研究者の交流会、2008年11月21日、柏

「HapMap ハプロタイプの問題点と Asian haplotype の今後」 林 健志

また、確定ハプロタイプについて、林 健志が2006年9月に香港で行われた国際会議HGV2006の招待講演で発表し、その内容は「Nature Genetics」のミーティングレポート(Abecasis et al., Nature Genet. 39, 153-155, 2006)で取り上げられた。

D-HaploDBは、発現量的形質のマッピングに関する総説(Cookson et al., Nature Review Genetics 10, 184-194, 2009)でも紹介された。

転写開始点領域、および疾患感受性候補領域の解析により同定されたSNPおよびそのアレル頻度情報はdbQSNPデータベースに蓄積しているが、さらにそのほとんどをヒト多型情報の中心的データベースである米国NCBIのdbSNPにKYUGENのハンドル名で登録した(配列情報約8000件、アレル頻度情報約5000件)。これらについて国内外から実験条件の問い合わせなどの影響がある。

プールDNAを用いたゲノムワイド解析は海外でも報告があるが、我々のシステムは同じプールを独自開発技術である定量的SSCP法で解析する2次スクリーニングすることにより、効率的に疾患関連SNPを同定できるという特長がある。また、既知の関連領域も多数検出できていることから、アレイのマーカーがカバーしている範囲内では偽陰性も少ないと考えられる。この方法により、国内で初めてSLEのゲノムワイド関連解析を行い、疾患と関連の強い複数の領域の同定に成功した。

<達成できなかったこと、予想外の困難、その理由>

SLEに強く関連する領域を同定できたが、その関連の原因についてはまだ不明である。しかし、これはこれまで得られたゲノムワイド関連解析結果のほとんどに共通の問題であり、新たな生物学的課題を提起しているものと考えている。

また、既成のマイクロアレイを用いた共通の課題として、領域によりアレイに搭載されるマーカーの密度が偏っており、それにより検出にバイアスがかかっている点も、ゲノムワイド解析の結果を解釈する上で考慮しなければならない。

<今後の課題、展望>

プールDNAを用いた関連解析が非常に有用であることが、今回の研究で明らかになった。この方法を他の多因子疾患の解析にも役立てたいと考えている。その際には全胎状奇胎の解析により得られた日本人の確定ハプロタイプを考慮したうえでカスタムアレイを作製することにより検出力を高めることができるので、今後検討していきたい。また定量的SSCP解析は簡便で低コストの解析手法であるので広く多型解析に応用していきたい。

一方、超並列型次世代シーケンサーの普及に伴い、疾患の遺伝的要因でこれまでの解析では同定できなかったような稀な変異を大規模に同定することが可能となってきた。これまでのSLEの関連解析により同定されたSNPでは遺伝的要因の大部分は説明できないことより、マイクロアレイを用いた関連解析により同定するのが困難な稀な変異がSLE発症に関連している可能性が大きい。これを候補遺伝子の大規模シーケンス解析により解明することを目指す。このような解析にもプールDNAを用いることが有効であることが海外の研究で示されているので、これを取り入れて発展させる。これにより疾患感受性の原因となる変異を同定し、遺伝的要因の解明をめざす。また、疾患との関連の生物学的意義を検討し、これらにより疾患の予防・治療の手掛かりとなる知見を得る。

全胎状奇胎DNAについてはこれまでのすべてのタイピング結

果を統合し、CNVとSNPの連鎖不平衡構造を明らかにする。また、この情報を疾患の関連解析に利用する。さらに、超高速シーケンサーの利用が可能になってきたので、今後の展望として全ゲノムシーケンス解明を考えている。現状の技術ではヘテロ部位の読取りが困難であるので通常のディプロイドゲノムでは高深度で読み取ることが必要となりコストがかかる。CHMは全領域ホモであるのでシーケンスの解釈がディプロイドゲノムより容易であり、より少ない読み取り回数で全ゲノム配列を読み取ることができる。またこれまでのSNPタイピング結果があるので、解読結果の検証も可能である。これを利用して日本人ゲノムの多型の全体像を明らかにし、また今回の研究で同定された複雑なCNV構造がどのようにして形成されたかを配列切断点の読取りによって解明していきたいと考えている。

<研究期間の全成果公表リスト>

1) 論文/プロシーディング

1. 0910281337
Umeno J, Matsumoto T, Esaki M, Kukita Y, Tahira T, Yanaru-Fujisawa R, Nakamura S, Arima H, Hirahashi M, Hayashi K, Iida M. Impact of group IVA cytosolic phospholipase A (2) gene polymorphisms on phenotypic features of patients with familial adenomatous polyposis, *Int J Colorectal Dis*, in press.
2. 0906031619
Higasa K, Kukita Y, Kato K, Wake N, Tahira T, Hayashi K. Evaluation of haplotype inference using definitive haplotype data obtained from complete hydatidiform moles, and its significance for the analyses of positively selected regions. *PLoS Genetics*, 5(5): e1000468 (2009)
3. 0910282253
4. 0910282258
Kiyohara C, Washio M, Horiuchi T, Tada Y, Asami T, Ide S, Atsumi T, Kobashi G, Takahashi H, and the Kyushu Sapporo SLE (KYSS) Study Group. Cigarette smoking, STAT4 and TNFRSF1B polymorphisms, and systemic lupus erythematosus in a Japanese population. *J. Rheumatol.*, 36(10): 2195-2203 (2009)
5. 0910282251
Horiuchi T, Washio M, Kiyohara C, Tsukamoto H, Tada Y, Asami T, Ide S, Kobashi G, Takahashi H, and the Kyushu Sapporo SLE (KYSS) Study Group. Combination of TNF-RII, CYP1A1 and GSTM1 polymorphisms and the risk of Japanese SLE: findings from the KYSS study. *Rheumatology*, 48(9): 1045-1049, (2009)
6. 0910282256
Kiyohara C, Washio M, Horiuchi T, Tada Y, Asami T, Ide S, Takahashi H, Kobashi G, and the Kyushu Sapporo SLE (KYSS) Study Group. Cigarette smoking, N-acetyltransferase 2 polymorphisms and systemic lupus erythematosus in a Japanese population. *Lupus*, 18(7): 630-638 (2009)
7. 0811061612
Kozawa M, Kondo H, Tahira T, Hayashi K, Uchio E. Novel mutation in PAX3 gene in Waardenburg syndrome accompanied by unilateral macular degeneration. *Eye*, 23(7):1619-1621 (2009)
8. 0811061603
Kuga D, Mizoguchi M, Guan Y, Hata N, Yoshimoto K, Shono T, Suzuki SO, Kukita Y, Tahira T, Nagata S, Sasaki T, Hayashi K. Prevalence of copy number neutral LOH in glioblastomas revealed by genome-wide analysis of laser-microdissected tissues. *Neuro Oncol.* 10(6): 995-1003 (2008)
9. 0710182010
Miyagawa H, Yamai M, Sakaguchi D, Kiyohara C, Tsukamoto H, Kimoto Y, Nakamura T, Lee J-H, Tsai C-Y, Chiang B-L, Nagasawa K, Harada M, Tahira T, Hayashi K, Horiuchi T. Association of polymorphisms in complement component C3 gene with susceptibility to systemic lupus erythematosus. *Rheumatology* 47(2): 158-164 (2008)
- 10.0710182005
Guan Y, Hata N, Kuga D, Yoshimoto K, Mizoguchi M, Shono T, Suzuki S, Tahira T, Kukita Y, Higasa K, Yokoyama N, Nagata S, Iwaki T, Sasaki T, Hayashi K. Narrowing the regions of allelic losses of chromosome 1p36 in meningioma tissues by an improved SSCP analysis. *Int. J. Cancer* 122(8): 1820-1826 (2008)
- 11.0710182002
Qin M, Kondo H, Tahira T, Hayashi K. Moderate reduction of Norrin signaling activity associated with the causative missense mutations identified in patients with familial exudative vitreoretinopathy. *Human Genetics* 122(6): 615-623 (2008)
- 12.0901131327
- 13.0901131330
- 14.0901131334
Mitoma H, Horiuchi T, Tsukamoto H, Tamimoto Y, Kimoto Y, Uchino A, To K, Harashima S, Hatta N, Harada M, Mechanisms for cytotoxic effects of anti-TNF agents on transmembrane TNF-expressing cells: comparison among infliximab, etanercept and adalimumab, *Arthritis Rheum*, 58(5): 1248-1257 (2008)
- 15.0901131337
- 16.0901131341
- 17.0805271423
Tao K, Fujii M, Tsukumo SI, Maekawa Y, Kishihara K, Kimoto Y, Horiuchi T, Hisaeda H, Akira S, Kagami S, Yasutomo K., Genetic variations of Toll-like receptor 9 predispose to systemic lupus erythematosus in Japanese population, *Ann Rheum Dis*, 66, 905-909 (2007)
- 18.0801231543
Kondo H, Qin M, Tahira T, Hayashi K. Severe form of familial exudative vitreoretinopathy caused by homozygous R417Q mutation in frizzled-4 gene. *Ophthalmic Genetics* 28(4): 220-223 (2007).
- 19.0608071243
Mitsuyasu H, Kawasaki H, Ninomiya H, Kinukawa N, Yamanaka T, Tahira T, Stanton VP Jr, Springett GM, Hayashi K, Tashiro N, Kanba S. Genetic structure of the dopamine receptor D4 gene (DRD4) and lack of association with schizophrenia in Japanese patients. *J. Psychiatric Research* 41(9): 763-775 (2007)
- 20.0801231727
Yanaru-Fujisawa R, Matsumoto T, Kukita Y, Nakamura S, Yao T, Hayashi K, Iida M. Impact of Phospholipase A2 group IIa gene polymorphism on phenotypic features of patients with familial adenomatous polyposis. *Dis Colon Rectum.* 50(2): 223-231

- (2007)
- 21.0701291716
Kondo H, Qin M, Kusaka S, Tahira T, Hasebe H, Hayashi H, Uchio E, Hayashi K. Novel mutations in Norrie disease gene in Japanese patients with Norrie disease and familial exudative vitreoretinopathy. *Invest. Ophthalmol. Vis. Sci.* 48(3): 1276-1282 (2007)
- 22.0612191638
Higasa K, Miyatake K, Kukita Y, Tahira T, Hayashi K. D-HaploDB: a database of definitive haplotypes determined by genotyping complete hydatidiform mole samples. *Nucleic Acids Res.* 35: D685-689 (2007)
- 23.0701191537
Horiuchi T, Kiyohara C, Tsukamoto H, Okamura S, Nagasawa K, Harada M. et al. A functional M196R polymorphism of tumor necrosis factor receptor type 2 is associated with systemic lupus erythematosus: a case-control study and a meta-analysis. *Ann. Rheum. Dis.* 66(3): 320-324 (2007)
- 24.0608071156
Qin M, Kondo H, Uno H, Fujiwara E, Uchino E, Tahira T, Hayashi K. Novel OPA1 mutations identified in Japanese pedigrees with optic atrophy. *Molecular Vision* 12: 485-491 (2006)
- 25.0608071251
Tahira T, Okazaki Y, Miura K, Yoshinaga A, Masumoto K, Higasa K, Kukita Y, Hayashi K. QSNPlite, a software system for quantitative analysis of SNPs based on capillary array SSCP analysis. *Electrophoresis* 27(19): 3869-3878 (2006)
- 26.0608071232
Jespersgaard C, Larsen LA, Baba S, Kukita Y, Tahira T, Christiansen M, Vuust J, Hayashi K, Andersen PS. Optimization of capillary array electrophoresis single strand conformation polymorphism analysis for routine molecular diagnostics. *Electrophoresis* 27(19): 3816-3822 (2006)
- 27.0608071140
Higasa K, Hayashi K. Periodicity of SNP distribution around transcription start sites. *BMC Genomics* 7: 66 (2006)
- 28.0701291703
- 29.0701291653
- 30.0701291644
Horiuchi T, Morita C, Tsukamoto H, Mitoma H, Sawabe T, Harashina S-I, Kashiwagi Y, Okamura S. Increased expression of membrane TNF-alpha on activated peripheral CD8+ T cells in systemic lupus erythematosus. *Int. J. Mol. Med.* 17(5): 875-879 (2006)
- 31.0701291633
- 32.0601282002
Hata N, Yoshimoto K, Yokoyama N, Mizoguchi M, Shono T, Guan Y, Tahira T, Kukita Y, Higasa K, Nagata S, Iwaki T, Sasaki T, Hayashi K. Allelic Losses of Chromosome 10 in Glioma Tissues Detected by Quantitative Single-Strand Conformation Polymorphism Analysis. *Clin. Chem.* 52(3): 370-378 (2006)
- 33.0608071223
Hata N, Shono T, Yoshimoto K, Yokoyama N, Kawamura T, Nagata S, Matsumoto K, Hayashi K., An astroblastoma case associated with loss of heterozygosity on chromosome 9p. *J. Neurooncol.* 80(1): 69-73 (2006)
- 34.0601281838
- 35.0601281927
- 36.0601281855
Horiuchi T, Gondo H, Miyagawa H, Otsuka J, Inaba S, Nagafuji K, Takase K, Tsukamoto H, Koyama T, Mitoma H, Tamimoto Y, Miyagi Y, Tahira T, Hayashi K, Hashimura C, Okamura S, Harada M. Association of MBL gene polymorphisms with major bacterial infection in patients treated with high-dose chemotherapy and autologous PBSCT. *Gene Immun.* 6(2): 162-166 (2005)
- 37.0601281944
- 38.0601281938
Shikata K, Kukita Y, Matsumoto T, Esaki M, Yao T, Mochizuki Y, Hayashi K, Iida M. Gastric juvenile polyposis associated with germline SMAD4 mutation. *Am J Med Genet A.* 134(3): 326-329 (2005)
- 39.0601281903
Koyama T, Tsukamoto H, Miyagi Y, Himeji D, Otsuka J, Miyagawa H, Harada M, Horiuchi T. Raised serum APRIL levels in patients with systemic lupus erythematosus. *Ann. Rheum. Dis.*, 64(7): 1065-1067 (2005)
- 40.0601281908
Qin M, Hayashi H, Oshima K, Tahira T, Hayashi K, Kondo H. Complexity of the genotype-phenotype correlation in familial exudative vitreoretinopathy with mutations in the LRP5 and/or FZD4 genes. *Human Mutation* 26(2): 104-112 (2005)
- 41.0601281914
Tahira T, Baba S, Higasa K, Kukita Y, Suzuki Y, Sugano S, Hayashi K. dbQSNP: a database of SNPs in human promoter regions with allele frequency information determined by single-strand conformation polymorphism- based method. *Human Mutation* 26(2): 69-77 (2005)
- 42.0601281916
Kukita Y, Miyatake K, Stokowski R, Hinds D, Higasa K, Wake N, Hirakawa T, Kato H, Matsuda T, Pant K, Cox D, Tahira T, Hayashi K. Genome-wide definitive haplotypes determined using a collection of complete hydatidiform moles. *Genome Res.* 15(11): 1511-1518 (2005)
- 2) 学会発表
1. 0912032158
Tahira T, Masumoto K, Kukita Y, Okazaki Y, Yoshinaga A, Higasa K, Horiuchi T, Hayashi K. Identification of loci for systemic lupus erythematosus by pooling-based genome-wide association study. 57th Annual Meeting of the American Society of Human Genetics, 2008/11/12, Philadelphia, U.S.A.
2. 0911261457
Hayashi K. D-haplo: a genome-wide definitive haplotype determined using complete hydatidiform moles. The 8th international meeting on human genome variation and complex genome analysis (HGV2006), 2006/09/14, Hong Kong, SAR, China (招待講演)
- その他多数
- 3) 図書
1. 0910281323
Tahira T, Kukita Y, Higasa K, Okazaki Y, Yoshinaga A, Hayashi

K. Humana Press Inc. “Estimation of SNP allele frequencies by SSCP analysis of pooled DNA” in Single Nucleotide Polymorphisms, Komar AA ed. (2009) pp193-207.

2. 0612191648

林 健志, 羊土社, “メンデルから DNA を経てゲノムへ” バイオ研究マスターシリーズ「遺伝子工学集中マスター」, (2006) pp14-30.

4) データベース/ソフトウェア

1. 0602202030 (データベース)

「dbQSNP public」, ゲノムの遺伝子発現制御領域に存在する SNP を検索し, そのアレル頻度を日本人集団および西欧人集団で決定した結果を表示

<http://qsnp.gen.kyushu-u.ac.jp>

2. 0602202024 (データベース)

「D-Haplo-DB」, 全胎状奇胎を用いて直接決定した日本人ゲノムのハプロタイプおよび連鎖不平衡構造をゲノムブラウザー上に表示

<http://orca.gen.kyushu-u.ac.jp>

3. 0612191657 (ソフトウェア)

「QSNPlite」, シークエンシング及び DNA プールの PLACE-SSCP 法を行い, それらの結果を統合して SNP を同定し, 且つアレル頻度を決定するための Windows 上で稼動する解析ソフトウェア

<http://qsnp.gen.kyushu-u.ac.jp/placeSSCP/qsnp-lite/>

5) 研究成果による産業財産権の出願・取得状況

1. 0911261435

特許第 4009481 号「DNA 塩基配列における SNPs 検出法」発明者: 三浦 健一, 林 健志, 秋山 純子, 杉森 かおる, 特許権者: 富士通 (株), 林 健志, 発行: 平成 19 年 11 月 14 日

6) 新聞発表, その他顕著なもの

なし