

論文の内容の要旨

論文題目 ゲノムと組織特異的発現情報の統合で明らかにするマイクロ RNA の疾患病態
氏名 坂上 沙央里

<序文>

2001 年にヒトゲノムプロジェクトが終了し、個人の多様性を特徴づける一塩基多型 (SNP; single nucleotide polymorphism) の情報がマイクロアレイ技術の成熟により大量に正確に得られるようになった。ゲノム情報が正確に安価に大量に得られることを背景とし、2002 年に世界で初めてのゲノムワイド関連解析研究 (GWAS; genome-wide association study) が行われて以来、4000 件以上もの GWAS がこれまでに行われてきた。しかしながら関連 SNP がどのように生物学的な経路をたどってヒト疾患を形成するかは理解は道半ばなのが現状である。疾患病態の正確な理解には、①非コード領域の SNP ②統計学的有意水準を満たさない効果サイズの小さい SNP に関しても機能的な解釈を行うことが重要である。その際には、先述した個々の細胞上の基本的に不可変・一次元的なゲノム情報が、空間的分布・外部組織からの刺激・環境因子への暴露により、いつ、どのような組織で、どのような刺激を受けて遺伝子発現量を変化させ疾患発症にいたるのかという文脈を理解する必要がある。このようなこれまでのゲノム研究の潮流から、新たなオミクス層としてマイクロ RNA (miRNA) の網羅的発現情報とゲノム解析との統合解析することとした。本研究において、大規模疾患ゲノム解析結果・miRNA-標的遺伝子ネットワーク・組織特異的 miRNA 発現測定カタログの 3 つの情報を計算機上で統合解析することによる網羅的 miRNA インシリコ・スクリーニング手法 (MIGWAS ; miRNA enrichment analysis in GWAS) を開発し、ゲノム情報のみでは分からなかった miRNA の疾患病態への組織特異的文脈での関与の理解と、バイオマーカー・創薬対象候補 miRNA の同定を目的とした。

<方法>

MIGWAS では、ある miRNA が疾患発症に関与しているならば、その miRNA 自体の発現を調整する SNP にも、その miRNA が標的として翻訳後調節を行うターゲット遺伝子の発現を調整する SNP にも、疾患発症との関連が認められるだろうという仮説をおいた。これに基づき、miRNA と遺伝子の標的関係の中に GWAS の有意なシグナルが豊富に存在しているかどうか、そしてそれが特定の組織に特に集積しているかを帰無仮説

と比較し統計学的に評価する手法を開発した。本手法では、ゲノムワイドでの統計学的有意水準である P 値 $< 5 \times 10^{-8}$ を超えない関連を持つ疾患関連 miRNA-遺伝子をペアワイズに網羅的に検出することができ、実験的検証の候補を見出すために有用である。本手法では、まず初めに GWAS の要約統計量（過去の GWAS 研究で判明した SNP と疾患発症との関連の強さを P 値などの統計量で表し公開したもの）から得られる SNP と疾患との関連の強さの P 値を、連鎖不平衡関係等を考慮した遺伝子・miRNA レベルの P 値へと変換した。その後、miRNA-遺伝子標的関係情報を複数の予測アルゴリズム（miRDB、MiRmap、PITA、TargetScan）より得て、その中である閾値を超える結合関係のペアを同定した。このペアの中で、遺伝子・miRNA 共にその形質に対して有意な関連を示したペアの割合を求めた。この割合が統計学的に有意に高いかどうかの判断を行う必要があるが、この統計量の理論的分布は未知である。そこで、permutation 検定の概念に基づき、GWAS 結果から変換した遺伝子・miRNA の P 値を、共に毎回ランダムに入れ替えて同様の解析を多数回シミュレーションを行うことで、「miRNA-遺伝子ネットワークが疾患に関与していない」という帰無仮説下の統計量の分布を経験的に算出した。さらに我々は、どの組織で miRNA の働きが重要かの特定にも、MIGWAS の手法を応用した。まず FANTOM5 コンソーシアムが作成した 179 種類のヒト細胞での miRNA の網羅的な発現情報を再解析し、各細胞に特異的発現をしている miRNA を同定した。これを元に、GWAS の統計量から算出されたシグナルがどの細胞における特徴的 miRNA セットおよびそのターゲット遺伝子セットの発現調整領域に集積しているかの評価を行った。

さらに、関節リウマチを例に取り、疾患・対象研究に行い、日本人の早期関節リウマチ (RA) 患者 30 名と、日本人健常人参加者 33 名を対象とし Differential expression 解析を行うことで MIGWAS から同定された疾患関連 miRNA が実際に患者群で発現変動しているかを調べた。ここで挙げた候補 miRNA に関しては、新規 GWAS のメタ解析を行うことでインシリコレプリケーション研究を行った。最後に、候補 miRNA について eQTL データベースとの統合解析を行った。

<結果>

過去に世界で行われた延べ 350 万人以上を対象とした 49 の GWAS 結果を MIGWAS で解析した。身長・関節リウマチ・2 型糖尿病が特に病態への miRNA の関与が重要であることが判明し、2 型糖尿病においては欧米人・日本人において共通して miRNA の

関与が示された。さらに興味深いことに、疾患ごとに miRNA が強く関与している組織は極めて多様性が高かった。バセドウ病、C 型肝炎、関節リウマチ、全身性エリテマトーデス、クローン病などの免疫疾患では免疫細胞が miRNA の活動の場であり、LDL コレステロール・心筋梗塞発症においては脂肪細胞が miRNA の活動の場であることが分かった。それぞれの疾患において有意に関連を認める miRNA のリストの抽出にも成功した。

我々は、このようにしてコンピューター上で同定した miRNA が実際に疾患病態に寄与していることを実証するために、関節リウマチを例にとり、①関節リウマチ患者と健常人とで該当 miRNA に発現量の差異を認めるかの実証実験を行うとともに、②新規 GWAS メタ解析を実行し、該当 miRNA に強いシグナルが再現されることを示した。30 名の早期関節リウマチ患者・33 名の健常人における網羅的 miRNA シークエンスを行い発現変動解析をしたところ、MIGWAS により同定した miRNA は偶然より 98 倍高い確率で発現変動 miRNA を予測していることが分かった。さらに、新たに約 8 万人を対象とした関節リウマチ GWAS のメタ解析を行うと、これらの miRNA はいわゆるゲノムワイド水準を超えて強く頑健に関節リウマチの発症に関わっていることも示された。加えて、組織特異的トランスクリプトームカタログである GTEx データベースとの統合解析により、疾患関連 miRNA は、免疫担当細胞の代用としての血液細胞において標的遺伝子の発現量変動を介して関節リウマチの発症に関わっていることも示唆された。

<考察>

我々は、大規模 GWAS データ、miRNA-標的遺伝子ネットワーク、網羅的細胞特異的 miRNA 発現データの 3 者を MIGWAS により統合解析し、組織特異的な miRNA の疾患病態への関与を示すことに成功した。MIGWAS を多様なヒト複雑形質に応用することで、miRNA の疾患病態への関与は非常に組織特異性が高いことがはじめて明らかになった。さらに、MIGWAS を応用することで、どの miRNA が疾患に関与しているかの情報も得ることができる。この情報については、関節リウマチ患者・健常人とでの発現量変動解析により実証実験を行うことでバリデーションされ、さらに新規 GWAS メタ解析を行うことでインシリコにもレプリケーションされた。新規 GWAS メタ解析では関節リウマチの新規関連 2 領域の同定に成功した。また大規模 eQTL データベースとの統合解析により、候補 miRNA が候補標的遺伝子を介してどのように関節リウマチの発症病態に関与しているかも推測することができた。この 20 年で GWAS に代表されるヒト

ゲノム情報の解析がもたらした生命現象への知見は大きいですが、その一方で未だその解釈に課題が残っていることも事実だ。ここ数年はハイスループットな実験系の発展が花盛りを迎えており、マルチオミクス情報は急速に蓄積されつつある。今回の研究で実施したような多層の生命情報の統合解析を行うことで、生まれ持ったゲノム情報が、どのようにしてその後の時空間的な文脈や環境因子への暴露によって我々の多様性を形作るかを理解すること、それによりゲノム研究を臨床現場にまで還元することに貢献したいと考えている。