

Development and Application of Computer Programs for Mapping Disease-Related Genes Using Polymorphic Markers 多型マーカーを利用した疾患遺伝子解析のため のコンピュータプログラムの開発と応用

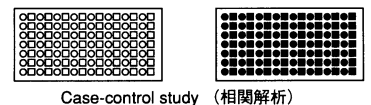
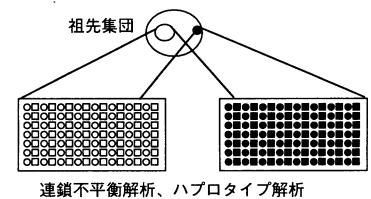
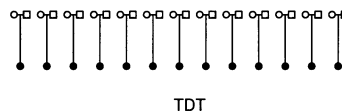
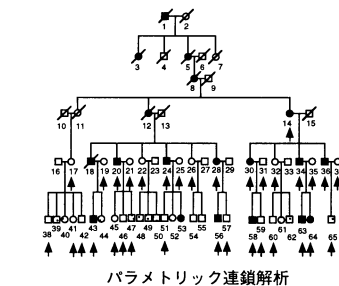
プロジェクトリーダー 鎌谷直之

東京女子医科大学

膠原病リウマチ痛風センター 教授

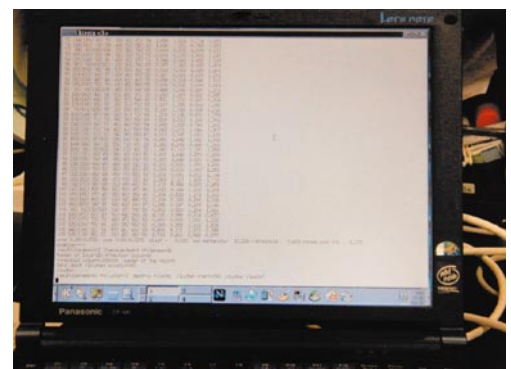


遺伝統計学の各種の手法



1. 研究の目的

ゲノム研究は、配列決定の時代から、社会的に役立つ、配列と形質（病気になるかどうかなど）との関係の解明の時代に入った。その目的を達成する手法として二つある。即ち、従来の分子生物学や配列のホモロジーサーチにより分子の機能を研究する事により形質との関係を解明する「知識に基づいたアプローチ」、そして個体の多型情報を用いる、新しい「遺伝統計学に基づいたアプローチ」である。この二つの方法が補いあって、これからのゲノム研究は進むと思われる。後者の手法は、個人のサンプルのデータ（遺伝子多型）と、表現型（病気かどうかなど）、さらには遺伝子地図情報を用い、遺伝統計学の論理で、大量のゲノム多型の情報とコンピュータプログラムを用いながら原因の場所をつきとめる方法である。本研究の目的の一つは遺伝統計的分析法のうち、既にほぼ完成している方法については、これを日本の研究者へ教育し、それを用いたゲノム解析を援助、あるいは請負い、疾患の原因遺伝子座の発見を手伝うことにより、本邦で形質に関係する遺伝子を発見することをサポートする事である。これからの重要な開発分野である連鎖不平衡解析、修正された相関解析については新たにプログラムを開発する。また、ゲノム研究では倫理的側面を十分考慮する必要がある。そのため、個人のサンプルと情報の匿名化の手法についても、新たにプログラムを開発する。



2. 研究の内容

1. 遺伝統計学分析支援体制の構築

統計学と遺伝学を利用して形質（病気になるかなど）に関係した配列を突き止める方法を「広義の連鎖解析」というが、本方法の理論と計算は難解であり、正しく理解し応用することは一般研究者にとって容易ではない。これらを日本の研究者が自由に利用できるように体制を整えるため、インターネット上で「遺伝統計学分析支援」ホームページを作る。それぞれの研究者が自分たちの得ることのできるサンプルの性質、目的とする疾患関連遺伝子の性質などから適当な手法を選ぶ事ができるようにする。さらに、そのために用いるべきプログラムとその正しい使い方をホームページ上で解説する。プログラムを使って得られた結果の解釈とまとめかたについても解説する。

2. 集団の構造化問題の大きさを測定し、構造に基づいた相関解析を行なうソフトウェアの開発

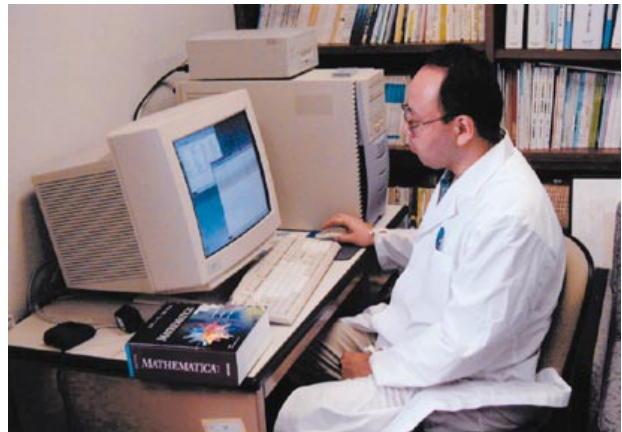
これまでに開発され、確立された遺伝統計的手法の多くはパラメトリック連鎖解析、ノンパラメトリック解析である。しかし、最も集めやすいヒトゲノム情報は家系ではなく個人のゲノム情報である。そして最もしばしば行なわれる病気の原因遺伝子の統計的研究はcase-control study、即ち相関解析である。しかし、相関解析は集団内に構造があると偽陽性に出ると言う根本的な欠陥がある。我々は集団に構造が存在するときに、個体の多数の多型情報を用いることによりこの構造の程度を計測したり、構造を持った集団からのサンプルでも偽陽性がでにくいようにして原因の遺伝子座を探すためのアルゴリズムを確立した。これをだれでも用いる事のできるソフトウェアとして整備、公開する。これを用いれば例えばインベダー法などで大量のSNP情報がcase-control studyで得られた場合に、それを偽陽性の出にくい形で分析することができる。プログラムを開発した後、実際にリウマチ性疾患などのサンプルを用いて、大量のゲノムデータを出し、そのデータをプログラムにかけてそれが有用であり、正しく原因関連遺伝子をつきとめるかどうかを確認する。

3. 連鎖不平衡解析を行なうソフトウェアの開発

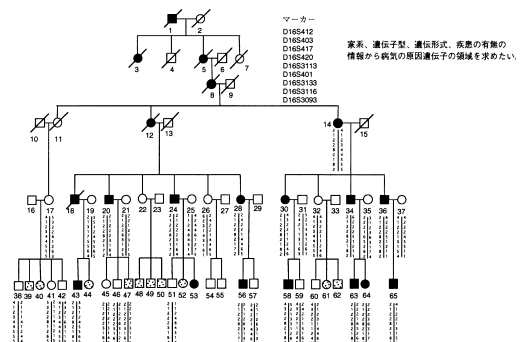
相関解析により病気などの原因がどこにあるかを解析して行くと、同一染色体上に連続して存在する多型部位についていずれも相関があると言う結果が出る事が多い。これは、単一の集団では、比較的頻度の高い疾患に関連した遺伝子の突然変異は共通のものが多く、その遺伝子と周辺の多型部位の間に連鎖不平衡が存在するためである（common disease-common variant）。このような場合、複雑な連鎖不平衡を整理し、疾患に本当に関係する遺伝子配列を発見する事を支援するソフトウェアを作成する。プログラムを開発した後、実際にAPRT欠損症などのサンプルを用いて、大量のゲノムデータを出し、そのデータをプログラムにかけてそれが有用であり、正しく原因関連遺伝子をつきとめるかどうかを確認する。

4. ゲノム倫理ガイドラインに準拠した匿名化のソフトウェアの開発

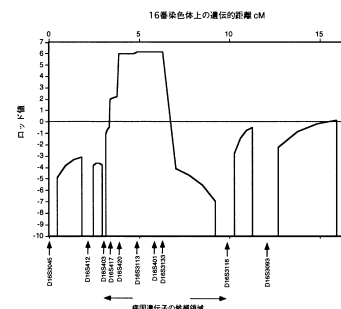
ゲノム解析には倫理問題を正しく理解し、個人情報が出たり、被検者が不安に陥らないように情報の匿名化を行なう必要がある。そのためには、遺伝子解析研究者に個人識別情報が伝達されないようにする必要がある。遺伝子解析のためのガイドラインでは、遺伝子解析に際しては個人識別情報の匿名化が必要であると明記している。しかし、匿名化のレベルにも種類があり、しかも表現型情報とサンプルの対応は間違えてはならない、などの困難な部分が多く存在する。我々のチームはこの匿名化のためのソフトウェアを開発し、多くの研究者が安心して用いることができるように整備する。



遺伝性痛風、腎不全、高血圧の家系と多型マーカーでのそれぞれの人々の遺伝子型



家系矛盾を解決 (incompcheck) したのち
パラメトリック連鎖解析 (Linkage package)
で多岐分析を行う。



3. 研究の体制

期 間：2000年6月～2005年3月

構 成：プロジェクトリーダー 鎌谷直之 東京女子医科大学
医学部 教授

コアメンバー1名

研究協力者11名

鎌谷直之をプロジェクトリーダーとして、小竹茂をコアメンバーとし、プロジェクトリーダー、コアメンバー以外の体制は次の通りである。山中寿、谷口敦夫、赤真秀人、小竹茂、針谷正祥、川口鎮司、市川奈緒美、斎藤仁之、北村豊、斎藤聡、氏名未定のポスドク2名

実施場所：東京女子医科大学・膠原病リウマチ痛風センター