

集団遺伝学に基づく疾患関連遺伝子検出理論の整備

● 舘田 英典

九州大学大学院理学研究院

〈研究の目的と進め方〉

疾患関連遺伝子のマッピングを効率的に行う方法の基礎理論を構築するために、疾患遺伝子頻度とマーカー遺伝子頻度の変化過程を解析して同時分布を求め、原因遺伝子検出可能性と種々パラメーターとの関係解明や効果的な遺伝子サンプリング法の検討を行う。計算機シミュレーションだけではなく、出来るだけ近似的手法を使って研究を進めていく。また実際のヒト集団でのDNA多型データを解析し、ヒト集団の遺伝的構造を探っていくと同時に、淘汰が働いたと考えられる機能的遺伝変異を探索する。

〈研究開始時の研究計画〉

遺伝的浮動の影響は受けるがより強い効果を持った遺伝子に連鎖した中立マーカー遺伝子の挙動を調べ、連鎖不平衡に関する近似的表現式を得ること、および集団遺伝学的手法によって弱い効果を持った遺伝的変異を検出するための理論の整備と、その実際集団への適用を目的に次の研究を計画した。

- 1) 任意交配集団で一遺伝子座に二つの対立遺伝子があり、ヘテロ接合体の適応度が両ホモ接合体の中間の値を取る場合(半優性)、連鎖(組み換え価 r)したマーカー遺伝子と有害(疾患関連)遺伝子(ホモ接合体で適応度を s 下げる)の頻度変化過程における同時頻度分布を解析する。遺伝子マッピングと対応させるため、ペネトランス等、遺伝子型同定可能性の問題を考慮に入れて、遺伝子サンプリングを行った時の標本分布を求める。同時に計算機シミュレーションを行い、近似解析で得られた結果の妥当性を検討する。
- 2) 優性の度合い(h)を任意に与えたときの有害突然変異遺伝子の頻度変化過程の性質を調べる。有害遺伝子頻度が低いところで条件付けるので、優性がない場合とよく似た結果が得られると考えられるが、計算機実験によりこの予測を検証する。
- 3) 集団サイズが変化した場合については、 D' の性質を調べ、平均値などモーメントに関する近似的解を求める。
- 4) 創始者効果があり集団サイズが減少した後指数的に増加した場合を仮定して、中立遺伝子の動態、中立遺伝子同士や、有害遺伝子と中立遺伝子の間の連鎖不平衡の動態を解析する。
- 5) 集団が複数に分かれて分集団構造を取り分集団間に移住(移住率 m)や融合がある場合として、単純な島モデルを仮定し、連鎖不平衡を解析する。異なる集団で作られた中立マーカーがどの程度他の集団で適用できるかについても、このモデルで検討する。
- 6) 複数遺伝子座での塩基配列多型データの提供を受けて、種々の中立性テストや連鎖不平衡に関する集団遺伝学的な解析を行い、日本人集団の遺伝的構造の平均的特性を推定するとともに、個々の遺伝子座で急速に頻度を変化させた遺伝的変異を同定し、common diseaseの原因候補である頻度の高い変異のうち、中立的ではない進化をしたものを探し出す。

〈研究期間の成果〉

1) - 3) については当初考えていた近似式による連鎖不平衡係数の計算値と、計算機シミュレーションによる計算結果が余りあわず、思ったように研究を進めることが出来なかった。4) - 6) については具体的に論文と対照させて成果を記述する。

4) については、まず集団サイズが確率的に変動する場合については、遺伝子の中立性を仮定して、拡散近似を使って平均ヘテロ接合頻度を求め、集団サイズの変動がどのように集団の有効な大きさに影響を及ぼすかを明らかにした(論文1)。集団の有効な大きさの定義にはいろいろあるが、ここでは平衡状態でのヘテロ接合頻度を使って定義した。この定義に従うと、従来考えられていたように集団の有効な大きさは調和平均となるのではなく、それより大きい値となることが示された。また算術平均よりは小さい値を取ることも示された。集団サイズの確率的变化が集団の有効な大きさにどのような効果を与えるかを具体的に見るために、集団サイズが二値の間を変動する場合について詳しい解析を行った。この場合、集団サイズの持続時間によって効果が異なり、早い変動(自己相関係数が小さい)では有効な大きさは調和平均に近づき、変動が遅くなると(自己相関係数が大きい)算術平均に近づくことが示された。調和平均と算術平均は場合によっては大きく異なるので、集団サイズの変動の早さによっては今まで考えられていた有効な大きさ(調和平均)と有効な大きさは大きく異なってくることがわかった。またこの単純な集団サイズが二値変動をするモデルでは、平均ヘテロ接合頻度だけでなくヘテロ接合頻度の分布も計算することが出来た。

ヒト集団を想定し、過去に集団サイズの変化(創始者効果など)が起こった場合について、中立マーカー遺伝子の変異から計算されるTajima's D (弱い自然淘汰やサイズ変化を検出する中立性テストとして良く使われている、実際このテストは6)の研究でも良く使用した)等の統計量の性質を明らかにした(論文5)。ここでは指数的に集団サイズが増加する場合について、遺伝子系図の形からこれらの統計量の平均・分散や分布を調べた。集団サイズの増加が急速な場合の極限では系図の形を求めることが出来、その系図の形を「急速増加極限系図」と名付けた。この系図の形から、上に述べた中立性テスト統計量の平均と分散を近似的に求めることができた。さらに、急速にサイズ増加が起こった時に近似的に従うと一般的に考えられている「星形系図」の場合については平均・分散以外に統計量の分布も求めることが出来た。「星形系図」では「急速増加極限系図」に較べて統計量の分布が極めて狭くなることがわかった。得られた統計量の平均・分散の式の形から、これらが増加期間の単調減少関数で有ることが示され、また R^2 統計量が集団サイズ変化の検出能力が高い事が知られているが、これはサイズの増加に対して、平均値が変化するよりもむしろ分散が小さくなることによることを示した。集団サイズがどの程度の早さで増加すると、急速増加極限近似が使える

ようになるかについてを調べるために、計算機実験による比較も行った。

5) については集団が分断化・融合を繰り返したとき、中立遺伝子の連鎖不平衡の動態を r^2 の平均値を指標に計算した(現在Theoretical Population Biologyに投稿中)。その結果このような分断化・融合を繰り返すことによって、染色体上の大きな領域にわたって連鎖不平衡が観測されることが明らかになった。染色体上での連鎖不平衡の広がりには集団が任意交配をしている集団とはかなり違ったパターンをとるので、適切に配置したマーカー間で連鎖不平衡を調べることによって、過去の集団構造を推定出来ることがわかった。島モデルの性質や異なる集団で作られた中立マーカーの適用可能性については検討出来なかった。

6) については、ヒトの複数集団においてFut2, Fut6, Pon1, AIM1の4遺伝子座での塩基配列多型データを解析した。Fut2遺伝子座では、ヨーロッパ人集団で頻度の高い変異が中立モデルから期待されるより有意に多く(Tajima's Dが有意に大きい)、この遺伝子座には何らかの平衡淘汰が働いていることが示唆された。この平衡淘汰は正常対立遺伝子とNull対立遺伝子の間に働いていると考えられる。またアジア集団はこの遺伝子座についてはヨーロッパやアフリカの集団から遺伝的に分化しており、このことからこの遺伝子座に淘汰が働いていることが推測される。一方Fut6遺伝子座ではこのような淘汰が働いたことを示唆する変異パターンは得られず、中立的な進化が起きていると考えられた(論文1)。

Fut2遺伝子座については他の人類集団からの遺伝的分化が進んでいると考えられるニューギニアの4集団での解析も行った(論文3)。この遺伝子座では上に述べたようにNull対立遺伝子が世界中の集団で高い頻度で存在しているが、ニューギニアではNull対立遺伝子の頻度が非常に低く、また正常対立遺伝子においても世界の他の集団では見られないSNPが見つかった。これらのことからニューギニアの集団が世界の他の集団から遺伝的に強く分化しており、人類進化の早いうちに隔離を受けたことが示唆された。しかしながらこの結果は一遺伝子座のデータのみに基づくものなので、今後より多くの遺伝子座での調査によって、この集団の遺伝的性質を明らかにする必要がある。

Pon1遺伝子座のプロモーター領域及びコーディング領域におけるSNPが循環器系疾患と関連があると言う報告がこれまでに幾つかなされているが、関連がないと言う報告も有り統一した見解が得られていない。そこで16.5kb離れたこれらの二つの領域、合わせて3 kbの塩基配列をアフリカ、ヨーロッパ、日本人集団の複数個体で決定し、多型のパターンについて解析した(論文4)。その結果、集団によってこの二つの領域間での連鎖不平衡量が異なっており、特に日本集団では2領域間で有意な連鎖不平衡を示すSNPの数が多いことがわかった。これらのことから、疾患とSNPの関連に関して必ずしも一致した結果が得られないのは、集団によってハプロタイプ頻度が異なることによることが示唆された。

AIM1遺伝子座では人間間で遺伝的分化が有る2ヶ所の非同義SNPが知られており、皮膚色との関連が指摘されてきた。そこでこのうちの一つのSNP(L374F)を含むエキソンとそのまわりの領域7.55kbについて複数個体で塩基配列を決定し、その多型を解析した(論文6)。その結果ヨーロッパ人集団で種々の中立性テスト統計量が有意になることが分かり、最近ヨーロッパ集団でselective sweep(変異の適応的固定による周辺連鎖部位

での多型量減少)が起こったこと示唆された。日光が弱くなることによる白人集団の適応の遺伝的基礎の一つと考えられた。

<国内外での成果の位置づけ>

ヒト集団での塩基配列決定を使った多型調査は国外では増えてきているが、国内では数少ない。このなかで適応的に進化をしていると考えられる変異が二つ見つかると、生物学的に重要な変異を検出する方法としての集団遺伝学的方法の有効性が示された。ランダムにサンプルされた遺伝子の塩基配列情報を使った中立性テストについては、計算機を使った研究が有るが、解析的近似法を使った研究はまだ殆どないので、有用な結果だと考える。

<達成できなかったこと、予想外の困難、その理由>

1) - 3)の研究において、最も簡単だと考えられたモデルでも、連鎖不平衡係数の二次以上のモーメントや分布を表す良い近似式を、期間中に得ることができなかった。疾患関連遺伝子が突然変異として生じた後、早い段階に起こる組み換えが連鎖不平衡係数に大きな影響を与えるが、この効果を未だ近似にうまく取り入れることができていないことが、満足の行く近似式を得ることができない一つの理由である。集団サイズが変化した場合については、同じ遺伝子座内の多型サイトの性質を調べる中立性テスト統計量の研究に時間がかかり、連鎖不平衡の研究を完成させることが出来なかった。また島モデルについても研究を進めることが出来なかった。これからの研究の重要課題である。自然淘汰を導入したモデルは予想以上に解析が難しく、有用な結果は得られなかった。これについてはやり方自体を検討し直す必要が有る。

<今後の課題>

集団サイズが変化する場合の連鎖不平衡については、必要な計算機は有るので、解析を進めて出来るだけ早く成果をまとめていきたい。また複数の遺伝子座で同じサンプルを使った塩基多型データが蓄積してきたので、これらをまとめて解析し、集団構造による効果と自然淘汰による効果を区別するような解析を進める必要が有る。

<研究期間の全成果公表リスト>

1) 論文/プロシーディング(査読付きのものに限る)

1. Y. Koda, H. Tachida, H. Pang, Y. Liu, M. Soejima, A. A. Ghaderi, O. Takenaka and Hi. Kimura Contrasting Patterns of Polymorphisms at ABO-Secretor Gene (FUT2) and Plasma a(1,3)Fucosyltransferase Gene (FUT6). *Genetics* 158, 747-756 (2001).

2. M. Iizuka, H. Tachida and H. Matsuda Neutral Model under Random Environments. *Genetics* 161, 381-388. (2002)

3. 0404081814
Y. Koda, T. Ishida, H. Tachida, B. Wang, H. Pang, M. Soejima, and H. Kimura DNA Sequence variation of the human ABO-Secretor locus (FUT2) in New Guinean populations: possible early human migration from Africa. *Human Genetics* 113, 534-541 (2003)

4 0404081843

Y. Koda, H. Tachida, M. Soejima, O. Takenaka and H. Kimura, Population differences in DNA sequence variation and linkage disequilibrium at the PON1 gene. *Ann Hum Genet.* 68(2),110-119 (2004)

5.

A. Sano and H. Tachida, Gene Genealogy and Properties of Test Statistics of Neutrality Under Population Growth. *Genetics* 169, 1687-1697 (2005).

6.

Mikiko Soejima¹, Hidenori Tachida, Takafumi Ishida, Akinori Sano¹, and Yoshiro Koda (2006) Evidence for recent positive selection at the human AIM1 locus in a European population. *Mol. Biol. Evol.* 23, 179-188.