



|              |  |
|--------------|--|
| Title        | Variation of gene-based SNPs and linkage disequilibrium patterns in the human genome |
| Author(s)    | 角田, 達彦   |
| Citation     | 大阪大学, 2007, 博士論文   |
| Version Type | VoR  |
| URL          | <a href="https://hdl.handle.net/11094/950">https://hdl.handle.net/11094/950</a>      |
| rights       |  |
| Note         |  |

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

|               |  |
|---------------|--|
| 氏 名           | つの だ たつ ひこ<br>角 田 達 彦  |
| 博士の専攻分野の名称    | 博 士 (医 学)  |
| 学 位 記 番 号     | 第 2 1 6 2 5 号  |
| 学 位 授 与 年 月 日 | 平成 19 年 11 月 22 日  |
| 学 位 授 与 の 要 件 | 学位規則第 4 条第 2 項該当   |
| 学 位 論 文 名     | Variation of gene-based SNPs and linkage disequilibrium patterns in the human genome<br>(ヒトゲノム中の遺伝子ベースの SNP と連鎖不平衡パタンのバリエーション) |
| 論 文 審 査 委 員   | (主査)<br>教 授 戸 田 達 史<br><br>(副査)<br>教 授 森 本 兼 曩 教 授 野 口 眞 三 郎   |

### 論 文 内 容 の 要 旨

#### [ 目 的 ]

ゲノムワイド関連解析を効率化するために、探索対象とする候補遺伝子を選出したり、マーカーを選定したりするには、ゲノム中の遺伝子ベースの一塩基多型 (SNP) および連鎖不平衡パターンに関する網羅的な情報を得ることが必要である。まず、全 SNP を解析するのは冗長で非効率的なので、ハプロタイプなどを最少の組合せで表現できる代表 (タグ) SNP の選定を行う。このとき、ハプロタイプブロック構造という、組換えホットスポットにはさまれた高い連鎖不平衡の領域を同定してから、その中でタグ SNP を選ぶことが提唱されているが、この構造すらも、ゲノム上の連鎖不平衡の度合いに大きく依存するので、連鎖不平衡のパターンをまず調べる必要がある。また、遺伝子内の SNP や連鎖不平衡の性質、ゲノム領域上の場所による連鎖不平衡の違い、組換え率との関係を調べることにより、遺伝子上あるいはゲノム上のどの領域・SNP を優先して探索するかを予測するための知識を得ることが必要である。これらを目的とし本研究を遂行した。

#### [ 方法ならびに成績 ]

対象とした SNP は、14,000 以上の遺伝子に対し系統的にリシーケンスすることによって得られた SNP セット (JSNP) である。それらを大規模な日本人集団 (564 人) に対し当研究センターでタイピングした結果を解析した。最終的に 74,842 SNP を解析に用いた。まず、遺伝子ベースのハプロタイプの性質に関する研究を行った。マイナーアレル頻度が 5% 以上の 65,080 SNP を用いたところ、頻度が 5% 以上の 46,558 のハプロタイプが、対象となった 13,419 遺伝子内に認められた。これらのハプロタイプを区別するタグ SNP を同定したところ、全部で 3 万 SNP 程度でよいことがわかった。補完計算の結果、常染色体上の遺伝子に対し日本人集団は 13 万のコモンハプロタイプ (頻度 > 0.05) を持ち、また必要となる SNP は 85,000 となるが、本研究ですでに 35% が捉えられたことがわかった。次に、遺伝子内およびゲノム上の位置の違いによる、SNP のアレル頻度や連鎖不平衡パターンの違いを調べた。まず、遺伝子内の SNP のアレル頻度を調べたところ、マイナーアレル頻度が 5% 未満のものは、非同義のエクソン SNP では 19.6%、同義のエクソン SNP で 14.4%、エクソン外 SNP で 12.0% というように、エクソン内の非同義 SNP も同義 SNP もエクソン外の SNP に比べアレル頻度が有意に低いことがわかった。このことは、浄化選択圧のかかり方の

違いであることが示唆される。また、同じ遺伝子内で、エクソン内の SNP どうしの連鎖不平衡が他の組合せと比べて平均して強いことがわかった。例えば、ほぼ完全連鎖不平衡にある SNP 対の頻度は、エクソン上の SNP どうしの方 (55%) が、他の組合せ (22%) に対し有意に多い。これらのエクソン SNP の対の 98% は、異なるエクソン内にあり、一つのエクソン内でなく遺伝子全体にわたる現象であった。これをロジスティック回帰分析を用い、距離、組換え率、アレル頻度の影響を除いて解析しても、エクソン上の SNP 対は他に比べ有意に連鎖不平衡が強いことがわかった。これも、浄化選択圧の現れである可能性がある。最後に、ゲノム領域上の連鎖不平衡について解析を行った。まず、連鎖不平衡は、別の実験で見積もられた組換え率とよく逆相関することがわかった。しかし、その関係はゲノム上の領域によってかなりばらつきがあり、相関がほとんど見られない領域もあることがわかった。6 番染色体短腕は、そのような相関が最もよく見られる領域の一つである。しかし同時に、その中の MHC 領域を含む 7 Mb の領域では、連鎖不平衡が低い領域をまたがって再び高くなるという奇妙なパターンも見つかった。そのような特徴的なパターンはこの領域のみに見られることもわかった。

#### [ 総 括 ]

本研究から、まず、遺伝子ベースのハプロタイプの本数は大変少なく保たれ、すなわち大人数の患者・対照者に対しタイピングすべきタグ SNP の数も大変少なくすみ、遺伝子ベースの関連解析は効率的であることがわかった。遺伝子内では、同義 SNP も非同義 SNP 同様、疾患関連 SNP として優先し探索することの意義づけが示唆された。またエクソン上の SNP 対の連鎖不平衡は、遺伝子内の他の SNP の組合せよりも強く、同様に選択圧がかかっていると考えられる。すなわち遺伝子内の 2 SNP 間の連鎖不平衡をくずすアレルの組合せが疾患の原因となりうることが示唆された。ゲノム上の連鎖不平衡が組換え率と明確な逆相関を示さない領域や、複雑なパターンを見せる MHC 領域などは、背景のメカニズムおよび疾患との関係が未知であり、将来より綿密な探索が行われるべき領域であることを示している。

#### 論文審査の結果の要旨

ゲノムワイド疾患関連遺伝子研究を効率化するには、SNP や連鎖不平衡の情報を網羅的に得、代表 (タグ) SNP の選定、優先探索すべき領域や SNP の予測が必要である。本研究では 564 人の日本人の 7 万 5 千 SNP のタイピングデータで 1 万 3 千以上の遺伝子のハプロタイプ、タグ SNP、連鎖不平衡のパターンを解析した。

- 1) 遺伝子内のコモンハプロタイプ 4 万 6 千とタグ SNP 3 万 (全数の 35% 以上) を同定し、遺伝子あたりそれぞれ 4 本、3 個程度と限られていることを見いだした。
- 2) 同義 SNP もエクソン外 SNP よりアレル頻度が低いこと、遺伝子内のエクソン SNP どうしの連鎖不平衡が他の組合せより強いこと、MHC 領域で独特な連鎖不平衡パターンがあることを発見した。

本研究では、日本人でのゲノムワイドな遺伝子 SNP・連鎖不平衡のパターンを初めて解析し、ハプロタイプやタグ SNP の同定、同義 SNP やエクソン SNP 対への選択圧の示唆、複雑な連鎖不平衡パターンの存在を発見し、ゲノムワイド関連解析による疾患遺伝子研究に貢献したので学位授与に値すると認める。