

平成21年 5月20日現在

研究種目：若手研究（B）  
 研究期間：2006～2008  
 課題番号：187101 59  
 研究課題名（和文） ヒト染色体間の相互作用情報の抽出と染色体の核内配置との関連性に関する研究  
 研究課題名（英文） Interaction among human chromosomes and its relationship to spatial organization of chromosomes in a nucleus  
 研究代表者 小柳 香奈子 (KOYANAGI KANAKO)  
 北海道大学・大学院情報科学研究科・准教授  
 研究者番号：20 362840

研究成果の概要：本研究の目的は、ヒトゲノム情報解析から染色体間の相互作用を示す情報を抽出し、核内における染色体の核内配置との関連性を検証することである。本研究では、染色体間の相互作用を示す情報として、ヒトゲノム中にみられるプロセス型偽遺伝子とゲノム内重複領域の染色体分布を解析し、染色体間の偏りを検出した。またこの研究の過程において、種特異的な重複領域を高精度に同定する目的で、分子進化解析に基づくゲノム直系領域の同定方法を開発した。

## 交付額

(金額単位：円)

|        | 直接経費      | 間接経費    | 合計        |
|--------|-----------|---------|-----------|
| 2006年度 | 252,325   | 0       | 252,325   |
| 2007年度 | 1,747,675 | 0       | 1,747,675 |
| 2008年度 | 700,000   | 210,000 | 910,000   |
| 年度     |           |         |           |
| 年度     |           |         |           |
| 総計     | 2,700,000 | 210,000 | 2,910,000 |

研究分野：複合新領域

科研費の分科・細目：ゲノム科学・基礎ゲノム科学

キーワード：遺伝子、ゲノム、進化、染色体、核内配置

## 1. 研究開始当初の背景

ヒトゲノムプロジェクトが終了し、ヒトゲノムの一次配列の全貌が明らかになった。現在このゲノムの機能を解明すべく、ゲノム上に存在する全遺伝子がいつでもどのように機能しているかといった発現解析や転写調節に関する解析がさかんに行われている。さらにここ数年、より高次のゲノム構造、すなわち染色体のクロマチン構造や核内配置が、遺伝子発現に重要な役割を果たしていることが明らかになってきている。とりわけ染

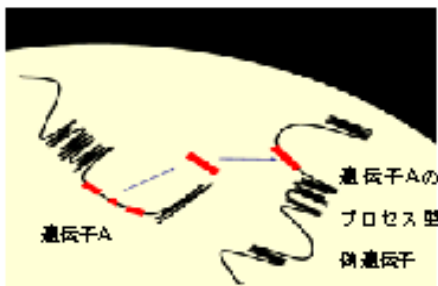
色体の核内配置は以下の三つの観点から注目に値する。(1) ゲノム機能の観点から、核内において複数の遺伝子がどのように協調して転写されるのか、また多細胞動物の発生・分化の過程で細胞種特異的な遺伝子発現がどのように確立・維持されるのか、といった問題に染色体の核内配置が関与しているのではないかと考えられている (Cremer *et al. Nature Rev. Genet.* 2001 2:292-301)。(2) 医学的観点から、がん細胞にみられる染色体の転座が、染色体の核内配置と関係していることが示唆されている (Parada *et al.*

*Trends Cell Biol.* 2002 12:425-32)。 (3) 進化学的な観点から、染色体再編成によるゲノム上の遺伝子位置の変化が、種間における遺伝子発現の違いをもたらしていることが示唆されている (Marques-Bonet *et al. Trends Genet.* 2004 20:524-9)。また、申請者は平成2004～2005年度の科学研究費若手研究B「ヒトゲノム構造にみられる遺伝子位置の偏りの意義とその進化過程に関する研究」において、共発現する遺伝子ペアが、ゲノム上で head-to-head の向きに隣り合う位置に再編成されて進化してきた、ということ明らかにしている (Koyanagi *et al. Gene* 2005 353:169-76)。このような進化における染色体再編成と遺伝子発現変化の関連性に関する研究は、これまで染色体の核内配置という観点からは捉えられてはならず、核内配置との関連から説明される可能性がある。

## 2. 研究の目的

本研究の目的は、ヒトゲノム情報解析から染色体間の相互作用を示す情報を抽出し、核内における染色体の核内配置との関連性を検証することにある。染色体の核内配置に関する研究は、ここ数年の間に実験解析からいくつもの進展があり、間期の核内には染色体テリトリーとよばれる領域が存在し個々の染色体が特定の区画を占有していることが示されている。しかしながら、染色体間の位置関係は決定的である (特定の染色体と染色体同士が常に近傍にある) という報告がある一方で、それらは動的かつランダムであるという矛盾した報告があったり (Bolzer *et al. PLoS Biology* 2005 3(5):e157, Parada *et al. Trends Cell Biol.* 2002 12:425-32 等)、顕微鏡観察に基づく実験解析からは染色体上の位置情報まで考慮した詳細な網羅的解析を行うことが難しい、といった問題がある。

そこで本研究では染色体間の相互作用を示す情報として、ヒトゲノム中にみられるプロセス型偽遺伝子の挿入のされ方の染色体間の偏りやゲノム内重複領域の染色体分布に着目し、ゲノム情報解析によるアプローチ



からこの問題の解明を試みる。例えば、ヒト

図1 プロセス型偽遺伝子

ゲノムにおいては、mRNA が逆転写されゲノム上の他の位置に挿入されたプロセス型偽遺伝子が約 8000 個存在するといわれている (Zhang *et al. Genome Res.* 2003 13:2541-58)。核内の染色体テリトリーの存在を考慮すると、このプロセス型偽遺伝子とその元となった親遺伝子は、核内で近傍に存在する染色体間に多くみられることが予想される (図1)。そこで、プロセス型偽遺伝子とその親遺伝子が存在する染色体の分布を調べ、統計的に有意に偏った分布をしている染色体群の情報を抽出する。同様にゲノム内重複領域の情報も利用し、それらの分布の染色体間における偏りを検出することで、染色体間の相互作用情報を網羅的に抽出する。抽出した情報と、既知の染色体核内配置に関する情報等を比較することで、これらの間の関連性を検証する。

## 3. 研究の方法

### (1) プロセス型偽遺伝子の同定とその染色体分布

プロセス型偽遺伝子とその親遺伝子は、エクソン数に注目することで区別・同定が可能である (プロセス型偽遺伝子はスプライシングを受けるのでエクソン数が1となる)。ヒトゲノム中に存在するプロセス型偽遺伝子とその親遺伝子について (Sakai *et al. Gene* 2007 389(2):196-203)、その染色体分布を調べ、統計的な偏りについて検証した。具体的には、各染色体ペアの分布がそれ以外の染色体ペアと比べて有意に異なるかについてカイ二乗検定を行った。

### (2) 重複領域の同定とその染色体分布

重複領域については、より精度の高い解析を行う目的で、材料をヒトとチンパンジーの種分岐後に起きた重複に絞って解析を行うことにした。これまでの解析から、チンパンジーゲノム概要配列データの質は低く、未決定または曖昧な状態である部分が存在することが明らかとなっている。また、現在公開されている BLASTZ 等を用いたゲノムアラインメントデータの多くは、重複領域についてのデータは存在しないか、存在しても重複の時期を考慮していないものが多く、網羅的なヒト系統特異的重複領域の同定がなされていない。そこで本研究では、チンパンジー BAC クローン配列 (3261 本、570,393,866 bp)、外群としてマカクゲノム (macaque genome NCBI build 1) を材料に用いて、ヒト系統特異的重複領域の同定を行った。具体的にはまず、チンパンジー BAC クローンを断片化したものを問い合わせ配列として、ヒトゲノムおよびマカクゲノムに対して BLASTN を実行し、相同領域を同定した。

次に、相同性が確認されたゲノム領域について MAFFT (Kato *et al. Nucleic Acids Res.* 2002 30(14):3059-66) を実行しマルチプルアラインメントを作成した。作成されたマルチプルアラインメントに基づき、マカクを外群として近隣結合法による分子系統樹を作成し、ヒトとチンパンジーの種分岐後にヒトの系統で特異的に重複したゲノム領域を同定した。さらに、ブートストラップ解析を行い、信頼性の高いヒト系統特異的重複領域を選別した。同定された重複領域についてその染色体分布を調べた。

#### 4. 研究成果

##### (1) プロセス型偽遺伝子の結果

プロセス型偽遺伝子とその親遺伝子の染色体分布の結果を図2に示す。ゲノム上に存在する因子の染色体分布から染色体間の関係をみる場合には、あやまって間接的な関係を抽出する可能性がある。例えば染色体分節重複により染色体A・B・Cにホモログな重複領域がみられた場合、染色体Aから染色体Bに重複がおき、その領域がさらに染色体Cに重複されたのか (A→B→C、AとCの関係は間接的である)、染色体Aからそれぞれ染色体B・Cに重複が起きたのか (A→BとA→C、BとCの関係は間接的である) といった情報を得ることは難しい。これに対しプロセス型偽遺伝子とその親遺伝子は、エクソン数に注目することで区別が可能であるため、直接的な関係にあるもののみを方向性の情報 (親遺伝子→プロセス型偽遺伝子) 付きで知ることが可能である。

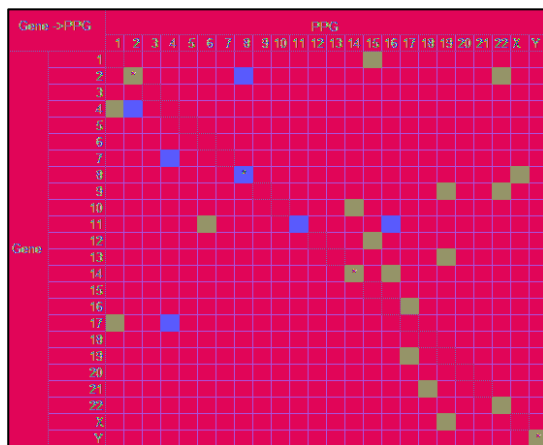


図2 プロセス型偽遺伝子(PPG)とその親遺伝子(Gene)の染色体分布

図2の赤いボックスは多くみられた染色体ペアを、青いボックスは少なくみられた染色体ペアを示す (星印は統計的に有意なペア)。この結果から、偽遺伝子とその親遺伝子

の染色体分布に偏りがあることが示された。この偏りが染色体の核内配置と関連しているのであれば、自分自身の染色体に多くみられることが予想される。しかしながらそのような結果は得られなかった。最近の研究から、プロセス型偽遺伝子の元となる mRNA が核外にできることが示唆されており (Pavlicek *et al. Trends in Genetics* 2006 22(2):69-73)、染色体分布に偏りはみられるものの、染色体の核内配置とは関連しないことが示唆された。

##### (2) 重複領域の結果

次に、ヒトゲノム内重複領域の染色体分布を調べた。ヒトゲノムの約5%はごく最近染色体分節重複を経験していることが知られており (Bailey *et al. Science* 2002 297:1003-7)、利用できる情報量が多いというメリットがある。ただし、上述の間接的な染色体間相互作用情報を抽出してしまうという問題がある。そこで、3箇所以上に分布するものは解析から除外することで問題を回避した。また、染色体の核内配置が進化の過程で変化し得ること、ヒトゲノムの塩基置換率が領域ごとに異なり得ること、の2点の理由から、ヒトとチンパンジーの種分岐後に起きた重複に絞り、進化時間を統一した材料を用いることにした。ヒトを含む霊長類の間では染色体の核内配置が保存していることが示唆されている (Tanabe *et al. Proc. Natl. Acad. Sci. USA* 2002 99:4424-9)。このために、ヒトゲノムとチンパンジーゲノムの高精度アラインメントを作成し、網羅的な分子進化解析を行い、チンパンジーとの種分岐後に重複したヒト系統特異的重複領域を同定した。

同定されたヒト系統特異的重複領域の染色体分布を図3に示す。解析に用いたチンパンジー-BAC クローンには染色体の偏りがありデータ量も十分とはいえないが、染色体分布に偏りのある可能性が示唆された。

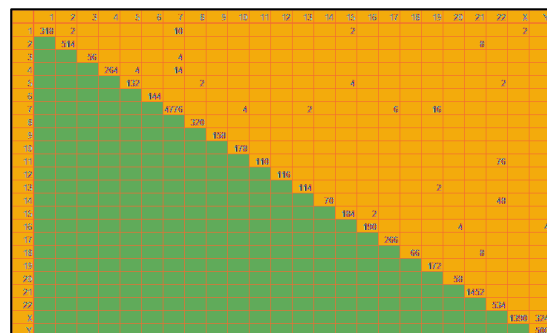


図3 重複領域の染色体分布

(3) まとめ

本研究により、染色体間相互作用情報が抽出され、プロセス型偽遺伝子やヒト系統特異的重複領域の染色体分布に偏りのあることが示唆された。これらの情報は、より詳細な実験解析を行うための有用な情報となり得る。今後精度の高いヒト近縁種のゲノムデータが蓄積すれば、より多くの精度の高い染色体間相互作用情報の抽出が可能となるだろう。

最新のゲノム研究から、本研究で用いたプロセス型偽遺伝子やゲノム重複領域以外にも、染色体間相互作用情報を含むと考えられるデータが急速に蓄積し始めている。例えば、癌細胞のゲノム変化に関する網羅的情報が蓄積し始めているが、がんゲノムみられる染色体再編成 (Campbell *et al.* *Nature genetics* 2008 40(6):722-729) も染色体核内配置の影響を受けている可能性がある。本研究で抽出された染色体分布の偏りとそのような染色体再編成の情報との比較を行えば、染色体核内配置と関連性の有無の検証が期待される。

また、本研究で開発した種特異的重複領域の同定方法は、重複領域を含むゲノム直系領域を高精度に同定することが可能である。したがってこの方法は、染色体の核内配置との関連性の研究に限らず、広く比較ゲノム研究に応用することが可能である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計3件)

- ① Matsuya A, Sakate R, Kawahara Y, Koyanagi KO, Sato Y, Fujii Y, Yamasaki C, Habara T, Nakaoka H, Todokoro F, Yamaguchi K, Endo T, Oota S, Makalowski W, Ikeo K, Suzuki Y, Hanada K, Hashimoto K, Hirai M, Iwama H, Saitou N, Hiraki AT, Jin L, Kaneko Y, Kanno M, Murakami K, Noda AO, Saichi N, Sanbonmatsu R, Suzuki M, Takeda J, Tanaka M, Gojobori T, Imanishi T, Itoh T. (2008)  
Evola: Ortholog database of all human genes in H-InvDB with manual curation of phylogenetic trees.  
*Nucleic Acids Research* 36: D787-792  
(査読有)
- ② Yamasaki C *et al.* (12 番目、他 137 名)

(2008)

The H-Invitational Database (H-InvDB), a comprehensive annotation resource for human genes and transcripts.

*Nucleic Acids Research* 36:D793-799  
(査読有)

- ③ Sakai H, Koyanagi KO, Imanishi T, Itoh T, Gojobori T. (2007)  
Frequent emergence and functional resurrection of processed pseudogenes in the human and mouse genomes.  
*Gene* 389(2):196-203 (査読有)

6. 研究組織

(1)研究代表者

小柳 香奈子 (KOYANAGI KANAKO)  
北海道大学・大学院情報科学研究科・准教授  
研究者番号：20362840

(2)研究分担者

なし

(3)連携研究者

なし