



東北大学

2013年 8月 5日

報道機関各位

東北大学大学院生命科学研究科

## ヒトゲノム上に遺伝子重複砂漠を発見 ～病気に関する遺伝子探索の新技术に期待～

### 概要

私たちは約2万の遺伝子を持っていますが、遺伝子が増えたり（遺伝子重複；注1）消失したりすることで、各々の遺伝子の数が個人によって違っていることがあります（コピー数多型；注2）。この遺伝子数の違いは、しばしば、自閉症や知的障害といった病気の原因となることが知られています。これまでに、コピー数多型のある遺伝子は、ゲノム中に偏って存在している事が報告されていますが、その原因についてはほとんど分かっていませんでした。東北大学大学院生命科学研究科生物多様性進化分野の牧野能士助教と河田雅圭教授はアイルランド・トリニティカレッジのイーファ・マックライザット博士と共同で、特定のタイプの遺伝子群が周辺に存在する遺伝子のコピー数多型を抑制していることを突き止めました。牧野助教らは、脊椎動物の初期進化で起きた全ゲノム重複に由来する遺伝子「オオノログ」（注3）に着目し、オオノログとその他の遺伝子のゲノム上の距離を調べました。その結果、オオノログの近くにある遺伝子はコピー数多型がない傾向にあり、逆にオオノログから離れて存在する遺伝子の多くがコピー数多型を示しました。特に、オオノログが高密度で存在するゲノム領域では、コピー数多型が強く抑制されていました。そのような領域は、数億年に渡る脊椎動物の進化過程において遺伝子が数を増やせない遺伝子重複砂漠であることも明らかにしました。今回の研究成果は、オオノログを含むゲノム領域の重複や消失が非常に有害（致死や病気を引き起こす）であることを示唆しています。このことから、遺伝子重複砂漠の領域内に存在しているコピー数多型は、病気と関係している可能性が高く、その領域のコピー数多型を調べることで、病気に関する遺伝子の効率的な探索が可能になると期待されます。本研究成果は、8月6日の英科学誌“*Nature Communications*（ネイチャー・コミュニケーションズ）”に掲載される予定です。

**(注1)遺伝子重複:** 一つの遺伝子がコピーされて二つの遺伝子になることを遺伝子重複といい、重複により生じた遺伝子を重複遺伝子と呼びます。多くの場合、遺伝子の重複は有害でも無害でもありませんが、最適な遺伝子数が厳密に決められた遺伝子においては、遺伝子重複が病気を引き起こすなどの悪影響を及ぼします。

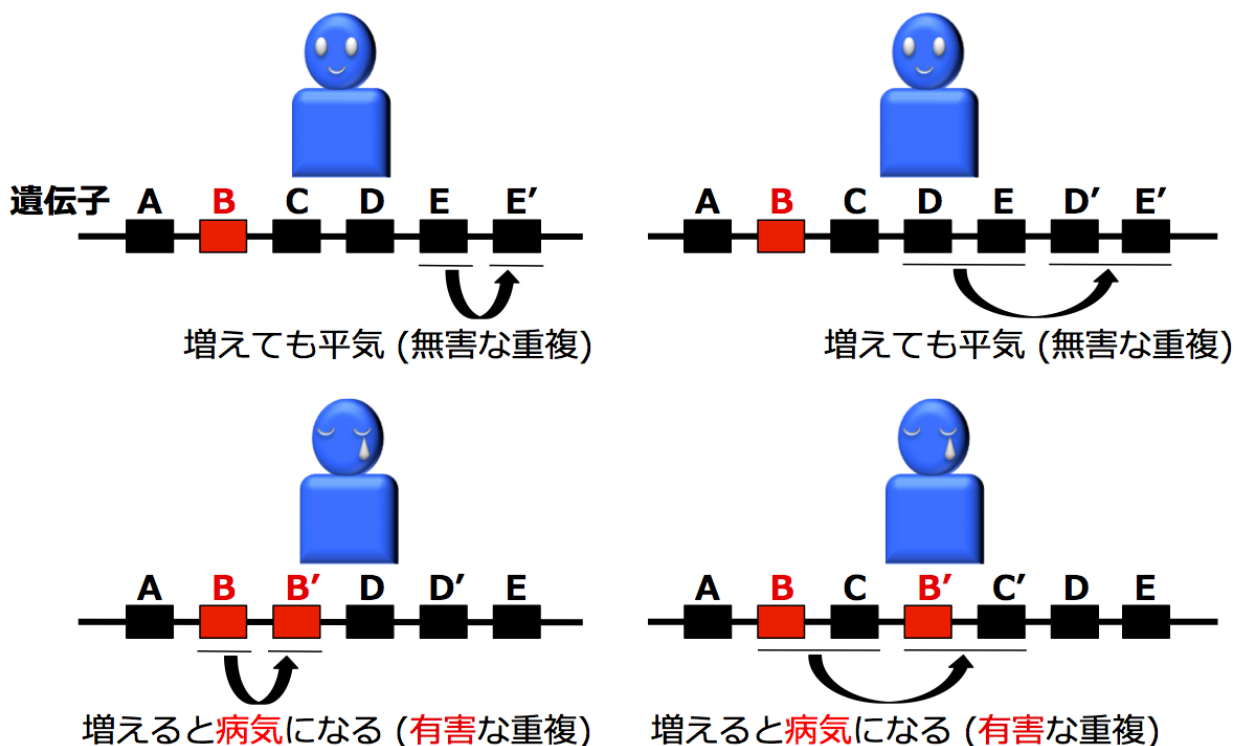
**(注2)コピー数多型:** 生物集団内における個体間比較において、ゲノム領域のコピー数の違いをコピー数多型と言います。ゲノム領域のコピー数は重複によって増加し、消失によって減少します。コピー数多型を示すゲノム領域中

に遺伝子が含まれる場合、遺伝子の数にも個体差が生まれます。例えば、ある人が遺伝子Aを1つ持つのに対して、別の人が遺伝子Aを2つ持つ場合、遺伝子Aはコピー数多型があると言えます。数多くのヒトのゲノムを調べた研究から、約30%のヒト遺伝子がコピー数多型を持つことが分かっています。コピー数多型を持つゲノム領域の長さは様々で、長いゲノム領域のコピー数多型では複数の遺伝子が含まれます。

**(注3)オオノログ:** 長い生物進化過程において、希にゲノム（全遺伝子）が重複する大イベントが起こることがあります。これを個別の遺伝子の重複（注1）と区別して全ゲノム重複と呼びます。ヒトを含む脊椎動物では、約5億年前の初期進化において2度の全ゲノム重複が起きたことが分かっています。全ゲノム重複により全ての遺伝子が倍加しますが、生じた2コピーの遺伝子は冗長であるため、多くの場合、1つが消失します。一方、全ゲノム重複後も消失せずに重複した遺伝子コピーを保持している遺伝子があり、このような重複遺伝子をオオノログと呼びます（※オオノログは全ゲノム重複による脊椎動物の進化を提唱された大野乾博士にちなんで名付けられました）。オオノログは最適な遺伝子数が厳密に決められている遺伝子群に多く、オオノログの重複や消失が有害な影響を与えるためにコピー数多型（注2）を持たない傾向にあります。オオノログとコピー数多型の関係についての詳細は先行研究のプレスリリースをご参照下さい。

「遺伝子進化パターンを調べるによりダウン症候群に関わる遺伝子を多数推定」

<http://www.tohoku.ac.jp/japanese/2010/05/press20100504-01.html>



**【図の説明】** 遺伝子B以外は増えても無害。遺伝子Bは増えることが有害で、重複すると病気や致死を引き起こす。そのため、遺伝子Bの重複遺伝子B'は子孫に伝わらない。遺伝子AとCは単独で増えた場合は無害だが、遺伝子Bの近くに存在するためにBとともに重複されることが多く、遺伝子AやCの重複遺伝子も子孫には伝わりにくい。このような背景から遺伝子Bの周辺では遺伝子が増えにくい遺伝子重複砂漠が形成される。

## 研究内容

### (背景)

私たちの体の設計図である遺伝子を構成するDNA配列には、個人個人で少しずつ違いがあります(塩基多型)。DNA配列の違いは、身体的特徴の違いや性格の違いなど様々な個性を生み出すのと同時に、病気の原因にもなり得ます。近年、ゲノム領域のコピー数が異なるコピー数多型(注2)の研究が進み、多くの遺伝子にコピー数多型があることが分かってきました。コピー数多型はゲノム領域の重複(注1)や消失により生じますが、多くの場合で有害でも無害でもありません。一方で、特定遺伝子のコピー数多型が自閉症や知的障害といった神経発達障害などの疾患の原因となることが報告されています。また、コピー数多型の分布はゲノム中で偏っていることが分かっていたのですが、その原因はよく分かっていませんでした。

コピー数多型の生じやすさは遺伝子によって異なっており、全ゲノム重複によって生じたオオノログ(注3)はコピー数多型を持たない傾向にあります。これはオオノログのコピー数の変動が有害であり、致死や病気を引き起こすためです。ゲノム領域のコピー数多型の中には複数の遺伝子が含まれることがあります(注2)。このため、オオノログと隣接する遺伝子では、オオノログとともにゲノム領域の重複(または消失)を経験する確率が高くなると推測されます(図1)。そこで我々は、オオノログと隣接する遺伝子は、たとえコピー数の増減が無害であっても、オオノログの重複や消失が成立しないためにコピー数多型を持ちにくいと考えました。

### (方法)

約2万あるヒトの遺伝子をオオノログとその他の遺伝子(非オオノログ)に分類し、全ての非オオノログについてゲノム上で最も近くに存在するオオノログまでの距離を求めました。また、コピー数多型データベース(Database of Genomic Variants)よりコピー数多型をもつヒトゲノム領域を抽出し、それらの領域に含まれる遺伝子をコピー数多型のある遺伝子としました。そして、オオノログまでの距離とコピー数多型のある非オオノログの関係を調べました。また、オオノログが高密度に存在しているゲノム上では、コピー数多型や遺伝子重複の頻度が少ないかどうか調査しました。

### (結果)

非オオノログにおけるオオノログまでの距離とコピー数多型の有無の関係について調べた結果、オオノログに隣接する非オオノログほどコピー数多型がない傾向にありました。そして、オオノログまでの距離とコピー数多型の存在には強い正の相関が観察されました(図2)。また、オオノログが高密度で分布するゲノム領域とコピー数多型の少ないゲノム領域(コピー数多型砂漠)に強い相関があることも分かりました(図3)。その傾向は特に、長いゲノム領域のコピー数多型で顕著でした(図3)。これは長いゲノム領域の重複ほどオオノログを含む可能性が高いためだと考えられます。今回の結果から、オオノログのゲノム上の位置が、ゲノム上のコピー数多型の分布を決定する要因であることが示されました。

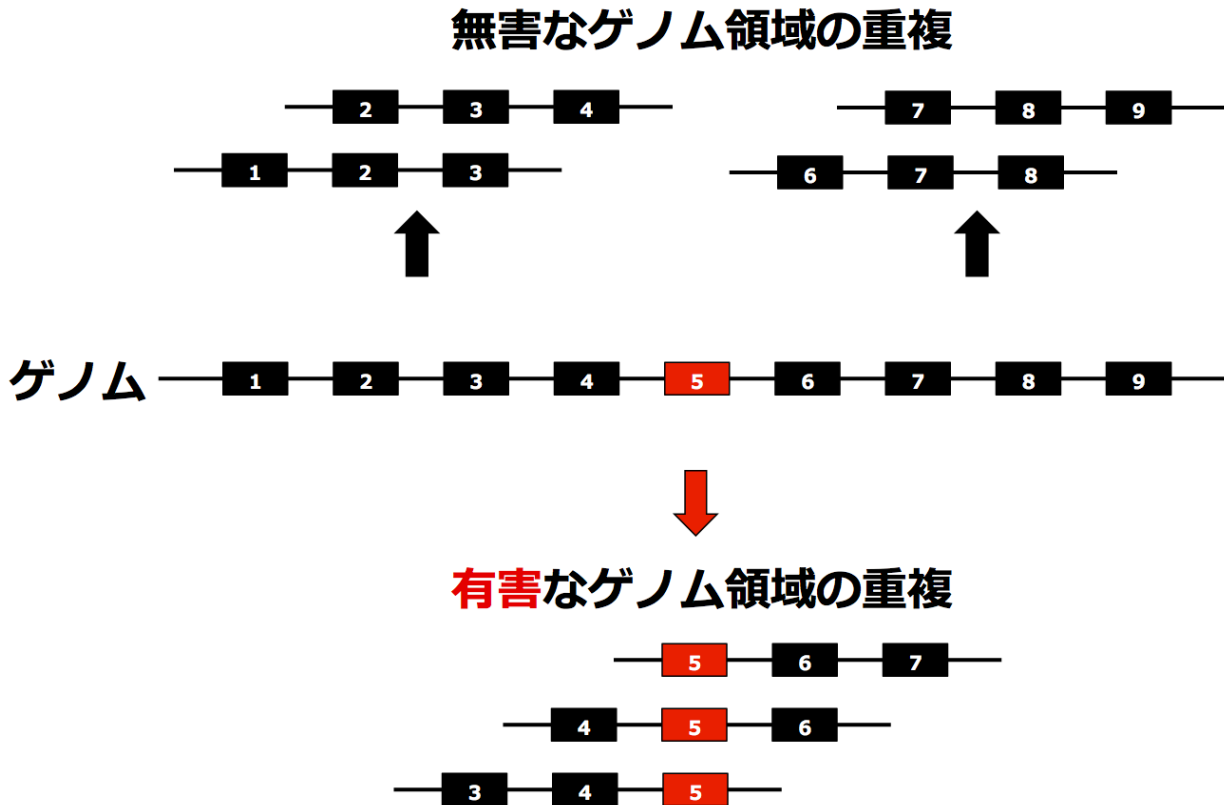
さらに8種の脊椎動物ゲノム(チンパンジー、マカクザル、マウス、ラット、イヌ、ウシ、オポッサム、ニワトリ)を用いた比較ゲノム解析により、オオノログが高密度に分布するコピー数多型砂漠は、数億年に渡る脊椎動物の進化過程においても遺伝子が重複しにくい **遺伝子重複砂漠**であることも明らかにしました。

オオノログとは対照的に、進化過程において重複を何度も繰り返して数を増加させている遺伝子ファミリーが存在します。特に臭いの感知に関する嗅覚受容体遺伝子は、ヒトにおいて数百、マウスでは約千も存在し、脊椎動物において最も大きな遺伝子ファミリーを形成しています。この遺伝子群のゲノム上の分布を調べたところ、ほとんどの嗅覚受容体遺伝子がオオノログから離れた領域に存在していました。このことは、ゲノム上のオオノログの位置が、他の遺伝子数の変動に大きく関わっていることを示唆しています。

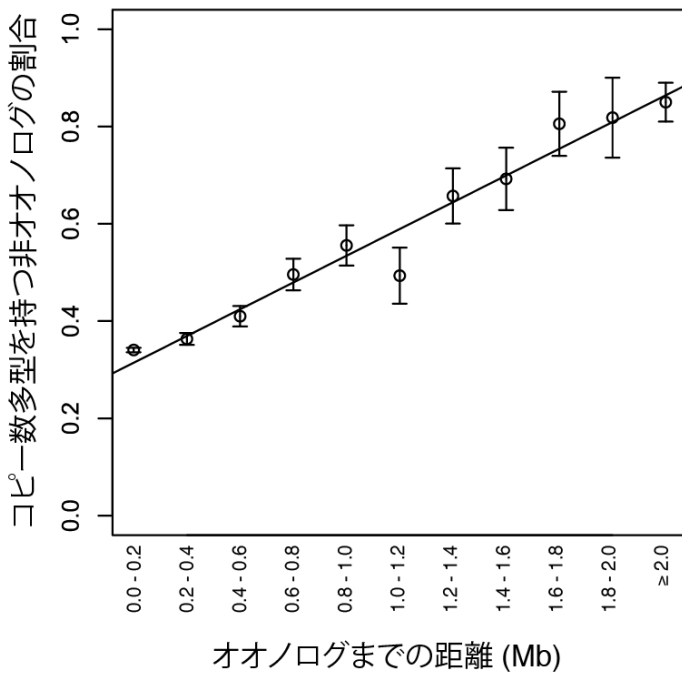
(今後の展望)

コピー数多型を網羅的に調べられるようになったのはごく最近のことですが、コピー数多型は遺伝的多様性を生み出す重要な変異源であるとして注目を集めています。我々は、こうしたコピー数多型の多くがオオノログから離れたゲノム領域で生じたことを示しました。一方、コピー数多型の研究が進むにつれて、コピー数多型が原因となる病気が多く報告されるようになりました。本研究において明らかにしたオオノログ周辺のゲノム領域ではコピー数多型が生じにくい知見を応用して、今後、オオノログ周辺のコピー数多型を調査することで、効率的な有害コピー数多型の同定が可能になり、病気との関連が疑われるコピー数多型の理解がさらに深まると期待されます。

図の説明

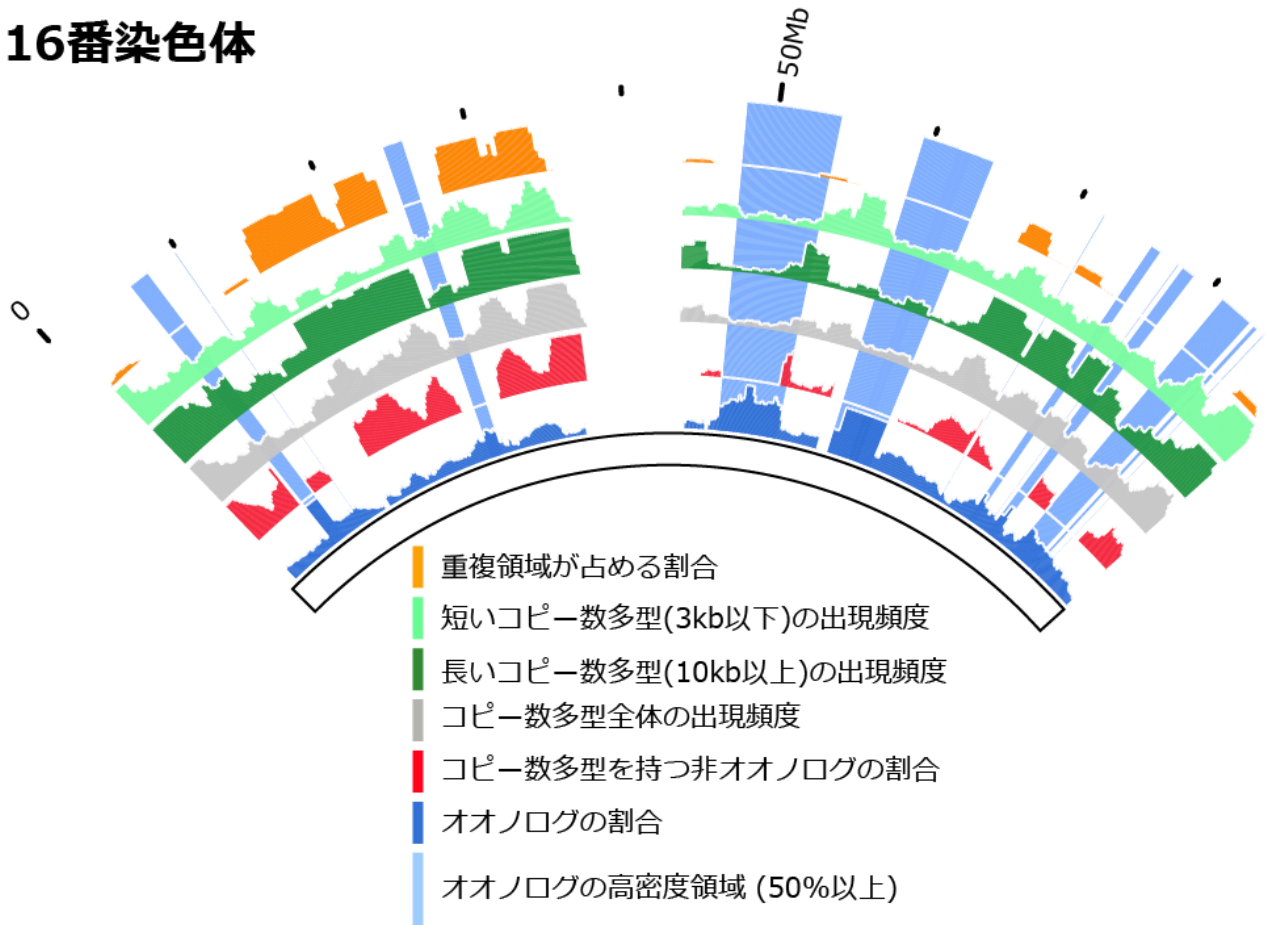


**図1. オオノログに近接する遺伝子のコピー数多型** 中央横線はゲノム、四角は遺伝子を示す。遺伝子のゲノム上の位置を左側から番号で示した（遺伝子1・9）。赤色四角はコピー数変動が有害なオオノログ、黒色四角はコピー数の変動が無害である遺伝子（非オオノログ）を表している。ゲノム上側と下側に示された短いゲノム断片は、遺伝子3つを含むゲノム領域が重複した場合のコピー数多型を模式的に示している。オオノログ（遺伝子5）を含まないゲノム領域の重複は無害であるが、オオノログを含むゲノム領域の重複は有害である。このため、オオノログと隣接する遺伝子3、4、6、7はコピー数多型が生じにくいと考えられる。



**図2. オオノログまでの距離とコピー数多型を持つ非オオノログの割合の関係** オオノログまでの距離をもとに非オオノログを11のグループに分類し（横軸）、それぞれのグループにおけるコピー数多型を持つ非オオノログの割合（縦軸）を調べた。オオノログとの距離が遠い非オオノログほどコピー数多型を持つ傾向が強い。

## 16番染色体



**図3. オオノログの密度とコピー数多型の関係(染色体16番の例)** オオノログ密度の高いゲノム領域(水色)では、コピー数多型をもつゲノム領域・遺伝子の頻度が低い。

本研究成果は(題目: Genome-wide deserts for copy number variation in vertebrates)、英科学誌 “*Nature Communications* (ネイチャー・コミュニケーションズ)” (日本時間8月6日) に掲載される予定です。

(お問い合わせ先)

東北大学大学院生命科学研究科

担当: 助教 牧野能士

電話番号: 022-795-6689

メール: tamakino@m.tohoku.ac.jp

ホームページ: <http://meme.biology.tohoku.ac.jp/klabo-wiki>