

令和 4 年 5 月 21 日現在

機関番号：12601

研究種目：基盤研究(B)（一般）

研究期間：2018～2020

課題番号：18H02514

研究課題名（和文）日本人の起源と進化 - 集団ゲノム学的アプローチによる縄文人由来ゲノム領域の抽出 -

研究課題名（英文）Population history of Japan

研究代表者

大橋 順 (Ohashi, Jun)

東京大学・大学院理学系研究科（理学部）・教授

研究者番号：80301141

交付決定額（研究期間全体）：（直接経費） 14,640,000 円

研究成果の概要（和文）：日本人男性345名のY染色体の全塩基配列決定を行い系統解析を行った。その結果、本土日本人男性では35.4%の頻度で観察されるが、他の東アジア人には観察されないクレードを発見した。遺伝子系図解析によって過去の人口変動を推定したところ、縄文時代晩期から弥生時代にかけて人口が急激に減少したことが示された。

47都道府県の全ゲノムSNPアリル頻度データを用いて解析を行い、都道府県間の遺伝的差異は、縄文人に由来するゲノム成分の程度と地理的位置関係によって説明できることを見出した。興味深い発見の一つは、近畿地方及び四国地方の人々が遺伝的に中国・漢民族に近いことであった。

研究成果の学術的意義や社会的意義

Y染色体の解析結果は、遺跡数や規模をもとに、縄文時代後期・晩期に縄文人の人口が減少したと推定した先行研究と符合するものであった。縄文時代晩期は世界的に寒冷化した時期であり、気温が下がったことで食料供給量が減ったことが、急激な人口減少の主な要因と思われる。47都道府県の遺伝的多様性解析の結果は、弥生時代以降に大陸から日本列島に移住してきた渡来人の中に、近畿地方と四国地方に進出・増加した人々が相対的に多かったことを示唆している。

本研究の成果は、縄文時代から弥生時代への移行期における日本列島人の進化史の理解に役立つと期待される。

研究成果の概要（英文）：A phylogenetic analysis for the entire Y chromosome of 345 Japanese males identified a clade that is observed at a frequency of 35.4% in mainland Japanese males, but not in other East Asians. The genealogical analysis indicated a sharp decline in population size of Japanese males from the late Jomon to the Yayoi period.

The analysis of genome-wide SNP allele frequency data of 47 Japanese prefectures revealed that genetic differences among prefectures can be explained by the extent of genomic components derived from the Jomon people and their geographic location. Of particular interest, it was found the modern Japanese people living in the Kinki and Shikoku regions were genetically closer to Han Chinese than people in the other regions were.

研究分野：ヒト進化遺伝学

キーワード：縄文人 渡来人 日本列島人 遺伝的多様性

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

これまでに、日本人のルーツを探るべく多くの研究が行われてきた。以前は、3つの説(交代説、混血説、連続説)が論じられていたが、現在は、日本列島にヒトが住み始めたのは約3万年前であり、約1万6千年前から縄文時代が始まり、約3千年前に大陸由来の集団と混血して弥生時代が始まったとする「混血説」が広く受け入れられている。埴原が提唱した二重構造仮説では、「現代日本人の祖先集団は東南アジア系で、旧石器時代から日本列島に住み縄文人となり、弥生時代から北アジア系(中国東北部など)の集団が日本列島に渡来し、在来の東南アジア系(縄文系)集団と徐々に混血し、日本人は今も二重構造を保っている。」と主張する。

われわれは、本土日本人、アイヌ人、琉球人、日本人以外のアジア人を対象に、HLA 遺伝子やゲノムワイド SNP (単塩基多型) を調べ、アイヌ人と琉球人は遺伝的に近縁であること、韓国人が本土日本人に最も近縁(渡来人の大部分は朝鮮半島由来)であることなどを報告してきた。遺伝学的研究によって多くの知見がもたらされたが、縄文人の純粋な子孫集団が存在しないため、縄文人と渡来人の混血過程を調べる遺伝学研究には限界があった。超並列シーケンサーを利用して縄文人 DNA を直接調べる試みも行われているが、日本の気候・土壌のせいで高品質の古代 DNA 試料を得ることは難しい。また数個体程度の解析では、縄文人の集団史や彼らの様々な形質の表現型頻度を知ることはできない。したがって、多数の縄文人・弥生人の高精度 DNA データが存在しない現状では、縄文時代から弥生時代にかけての日本列島人の進化史を調べるためには、従来とは異なるアプローチが必要である。

日本列島はユーラシア大陸東部に位置し、北海道、本州、四国、九州の四つの大きな島と多くの小さな島、そして琉球諸島で構成されており、南北約 3,000 km の弧状列島である。日本人集団は2つの小集団(アイヌと琉球人)と1つの大集団(本土人)から構成されている。これまでの頭蓋形態の研究から、アイヌと琉球人は本土人よりも縄文人に近く、東北地方の本土人はアイヌに近いことが示唆されている。遺伝学的研究からも、縄文人とアイヌおよび琉球人との関連性が確認されており、琉球人よりもアイヌの方が縄文人に遺伝的に近いことが示されている。

本土人と沖縄人(琉球人)との遺伝的分化を明らかにした先行研究は多くあるが、本土人の都道府県間の遺伝的差異は十分に理解されていない。日本は47の行政区画(都道府県)に分かれており、それらは9つの地方に分けられる。Yamaguchi-Kabata ら(2008)は、日本の7つの地方にある病院で治療された7,003人の日本人患者のゲノムワイド SNP データに基づき、日本人は遺伝的に分化していることを示した。しかし、四国地方と中国地方は研究に含まれていなかった。また、地方内の都道府県間の遺伝的差異も不明であった。したがって、本土人の遺伝的分化を理解するためには、47都道府県レベルでの高解像度な解析が必要である。

2. 研究の目的

本研究では、現代人のゲノムデータを用いて、(1)縄文人の集団サイズの変化、(2)渡来人との混血が現代日本人の遺伝的多様性に与えた影響を明らかにすることを目的とした。

3. 研究の方法

(1) 縄文人の集団サイズの変化

現代日本人は、縄文人と渡来人の混血によって誕生した集団の子孫である。組換えが起こるため、縄文人由来と渡来人由来の常染色体ゲノム領域を区別するのは困難であるが、組換えを受けない Y 染色体(領域)であれば、塩基配列の違いをもとに、縄文人由来の Y 染色体を同定できる可能性がある。そこで、日本人男性 345 名の Y 染色体の全塩基配列決定と系統解析を行った。次に、遺伝子系図理論に基づいて縄文人の集団サイズの変化を推定した。ここで、遺伝子系図解析と集団サイズの推定法の基礎理論について簡単に説明する。生物が子孫を残す際、親の遺伝子(DNA 配列)が複製され、いわゆる遺伝子コピーが子どもに伝わる。したがって、時間を遡って遺伝子コピーの親を順に辿っていくと、最終的に一つの共通祖先遺伝子(最も近い共通祖先遺伝子)に到達する。この過程を表現したものを遺伝子系図といい、遺伝子系図からその集団の人口変動を推測することができる。いま、N人の男性からなる集団があり、2人の男性から Y 染色体をサンプルしたとする。世代は離散的で、どの男性も次世代の任意の男性の親になる確率は等しい(どの男性も前世代の任意の男性が親である確率は等しい)と仮定する。男性集団のサイズは N で一定だとすると、サンプルした2人の共通

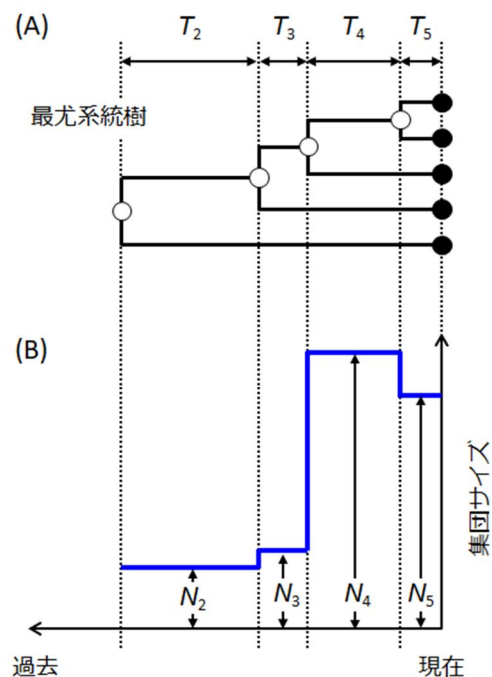


図1 5本のY染色体をサンプルした場合の最尤系統樹 (A) と集団サイズの推定値 (B)

率に等しい)と仮定する。男性集団のサイズは N で一定だとすると、サンプルした2人の共通

父親が前世代に存在する確率（これを合祖確率とよぶ）は $p_2=1/N$ であり、存在しない確率は $1-1/N$ である。すなわち、世代を遡って考えた場合、 k 世代前に初めて合祖する確率 c_k は $p_2(1-p_2)^{k-1}$ である。 $k c_k$ を $k=1$ から $k=N$ まで足し合わせると、初めて合祖するまでの期待世代数 N が得られる。この期待世代数 N は世代当りの合祖確率 p_2 の逆数である。同様の過程について、2本ではなく m ($m \geq 2$) 本で考えると、 m 本のうちの任意の2本が合祖して、 $m-1$ 本の独立な Y 染色体の系統となる世代当りの確率は、 m 本の中から任意の2本を選ぶ組合せ数 $m(m-1)/2$ を $1/N$ にかけて $p_m=m(m-1)/2N$ となる。一般的に、 N は十分に大きいので、3本以上の Y 染色体が同世代に合祖する確率は無視できる。世代当り、 p_m の確率で任意の2本が合祖し、 $1-p_m$ の確率で合祖しないので、 m 本の独立な Y 染色体の系統が $m-1$ 本の独立な系統となるのに要する期待世代数 T_m は、2本の場合と同様に考えて $1/p_m = 2N/\{m(m-1)\}$ である。したがって、 m 本の独立な Y 染色体の系統が $m-1$ 本の独立な系統となるまでの集団サイズ N_m は m と世代時間 T_m を用いて

$$N_m = \{m(1-m)T_m\}/2$$

と表すことができる。ここで、 m 本の独立な Y 染色体の系統が $m-1$ 本の独立な系統となるまでの集団サイズは一定と仮定している。

適当な分子進化モデルを仮定すれば、枝の長さが時間に比例した m 本の Y 染色体配列の最尤系統樹を得ることができる。最尤系統樹の枝長をもとに $m-1$ 個の T_m を求め、その値を上式に代入することで $m-1$ 個の N_m を推定することができる（図1）。図1では、合祖から次の合祖までの集団サイズを独立に計算したが、集団サイズが急に大きく変化するとは考えにくいので、 N_{m-1} は N_m に依存すると考えるべきである。われわれの研究では、その点についても考慮した、より洗練された統計手法を実装した BEAST 2.4.4 というソフトウェアを用いて、縄文人の集団サイズ変化を推定した。

(2) 渡来人との混血が現代日本人の遺伝的多様性に与えた影響

日本人集団の遺伝的構造を調べるために、47 都道府県に居住する日本人約 11,000 人のゲノムワイド塩基多型 (SNP) データを用い、日本人集団の遺伝的構造について調べた。ゲノムワイド SNP 遺伝子型は、次の二つのプラットフォームを用いて取得した: Illumina HumanCore 12 Custom BeadChip と HumanCore 24 Custom BeadChip。ここで、Hardy Weinberg 平衡 (HWE) P 値 < 0.01 、SNP コール率 < 0.01 、サンプルコール率 < 0.1 の SNP を選択した。1KG フェーズ 3 の 103 人の中国・漢民族 (CHB) の遺伝子型データと我々の 11,069 人の日本人のデータセットを併合したところ、二つのデータセット間で 138,688 か所の常染色体 SNP が共有されていた。日本人と CHB の全サンプルに対して主成分分析 (PCA) を行い、PCA プロットで CHB に近い 116 人の日本人サンプルを除外した。さらに、日本人サンプル中の近親者を除いた。最終的に、CHB を含む解析では 138,638 か所の常染色体 SNP を、日本人のみを対象とする解析では 183,708 か所の常染色体 SNP を用いることとした。

都道府県間の抽出誤差の差を調整するために、各都道府県から 50 個体を無作為に抽出した。都道府県レベルで日本人の遺伝的構造を理解するために、各都道府県から 50 個体を無作為に抽出して計算した頻度データをもとに PCA を行った。

4. 研究成果

(1) 縄文人の集団サイズの変化

345 名の Y 染色体の系統解析を行ったところ、日本人の Y 染色体は主要な 7 つの系統に分かれることが示された（図2）。さらに、韓国人・中国人を含む他の東アジア集団のデータを併せて解析した結果、日本人で 35.4% の頻度で見られる系統 1 は他の東アジア人に観察されないことが示された（図3）。系統 1 に属する日本人 Y 染色体の変異を詳細に解析したところ、系統 1 は YAP という特徴的な変異をもつ Y 染色体ハプログループ D1b に対

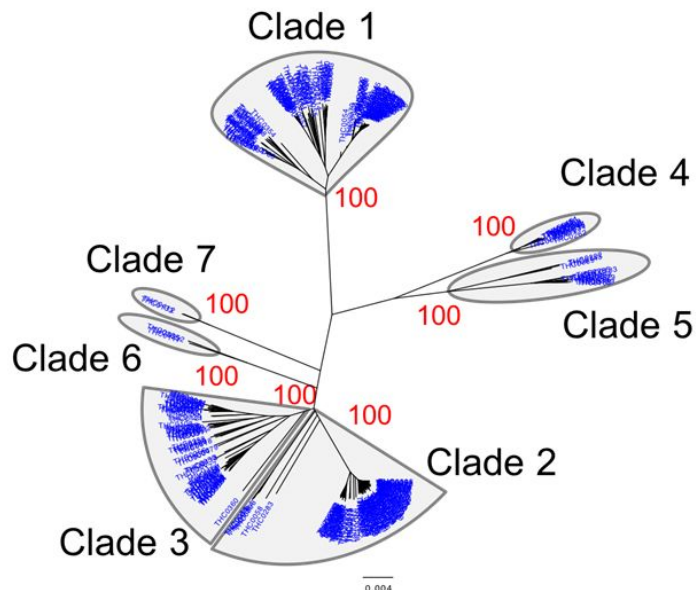


図2 日本人男性345名のY染色体系統樹

応していることが示された。YAP変異は、形態学的に縄文人と近縁と考えられているアイヌ人において80%以上という高い頻度で観察される。渡来系弥生人の主な母体である韓国人集団や中国人集団には系統1に属するY染色体が観察されなかった結果も踏まえると、系統1のY染色体は縄文人に由来すると結論できる。ちなみに、同一検体のミトコンドリアDNAの系統解析も行ったが、明らかに縄文人由来と想定されるような系統は存在しなかった。

縄文人一個体の常染色体ゲノムを解析した先行研究により、本土日本人のゲノム成分の約12%が縄文人に由来すると推定されている⁷⁾。このことは、縄文人に特異的、かつ固定していた(縄文人集団中での頻度が100%)変異であったとしても、混血後の本土日本人集団での頻度は12%程度にしかならないことを意味する。したがって、縄文人に由来する系統1のY染色体が35%以上の頻度で観察されることから、縄文人集団における系統1のY染色体頻度は相当高かったと思われる。すなわち、大部分の縄文人男性は系統1のY染色体を保有していたと考えられ、本土日本人がもつ系統1のY染色体の解析から、縄文時代の男性人口のサイズ変化を推定できると期待される。そこで、我々は系統1に含まれる122人のY染色体のみを対象に遺伝子系図(図4A)を推定し、前述の方法によって過去の集団サイズの変化を推定した。その結果、縄文時代晩期から弥生時代初期(およそ3,000~2,000年前)にかけて、人口が急減した後、急増したことが示された(図4B)。男性の数のみが変化したとは考えにくいので、女性も同様の変化を示したと思われる。縄文時代晩期は世界的に寒冷化した時期であり、気温が下がったことで食料供給量が減ったことが、急激な人口減少の要因の一つと考えられる。一方、弥生時代に入って人口が増加したのは、渡来系弥生人がもたらした水田稲作技術によって、安定した食料供給が可能になったからと思われる。

本成果は「混血説」にもう1つの示唆を与えてくれる。それは、弥生時代に起きた縄文人と渡来系弥生人の混血は平和的なものであったということである。現代本土日本人のゲノム成分の約12%が縄文人に由来すると推定されている。日本列島にはそれまで縄文人が住んでいたことを考えると、この数値は低すぎるように思われる。そのため、相当数の縄文人が渡来人との争いで殺された可能性が考えられる。もし縄文人と渡来系弥生人とのあいだで争いが起きれば、死ぬのは主に男性であろう。そうだとすると、最も残りにくい縄文人のゲノム成分はY染色体ということになる。しかし、常染色体のゲノム成分は約12%しか残っていないのに対し、本研究が示したように、縄文人に由来するY染色体の現代日本人中の頻度は35%を超えている。Y染色体だけがこれほど高い割合で縄文人に由来することを、Y染色体に作用した遺伝的浮動で説明するのは難しい。縄文人男性は渡来人によって殺されていないだけでなく、むしろ優位性を保ったまま混血したのではないだろうか。

本土日本人
他の東アジア人

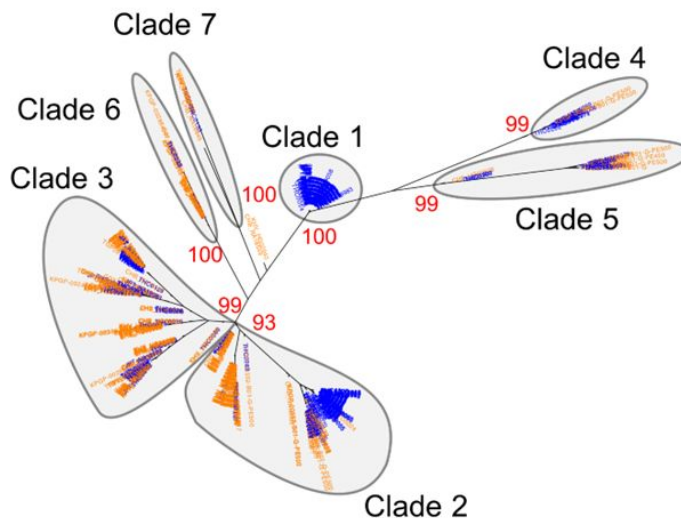


図3 日本人と他の東アジア人のY染色体系統樹

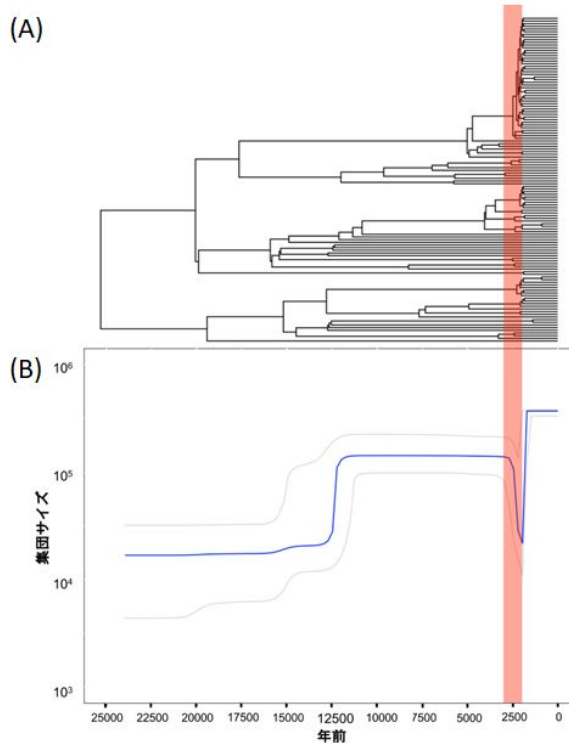


図4 122人の日本人Y染色体の遺伝子系図(A)と推定した集団サイズの変化(B)
(B)の青色の曲線が推定値であり、灰色の曲線は95%信用区間である。

(2) 渡来人との混血が現代日本人の遺伝的多様性に与えた影響

都道府県レベルで日本人の遺伝的構造を理解するためにPCAを行った(図5)。PC1とPC2の分散の割合は、それぞれ3.1%と2.6%であった。同一地方の都道府県は近接していた。PC2は、都道府県庁所在地の緯度(ピアソンの相関係数 $R=0.81$ 、 P 値= 3.21×10^{-12})および経度(ピアソンの相関係数 $R=0.85$ 、 P 値= 2.38×10^{-14})と強く相関していたが、PC1は相関していなかった。したがって、PC2は各都道府県のおおよその地理的位置を反映しており、これは隣接する都道府県間の移住や混血によるものと考えられる。次に、PC1に着目すると、沖縄は他の都道府県から離れており、東北地方と九州地方の県は沖縄に比較的近い位置にあった。沖縄から最も遠いのは近畿地方と四国地方の県であった。沖縄県は琉球人の遺伝的特徴を反映しており、琉球人が縄文人に遺伝的に近いことをふまえると、PC1は縄文人との遺伝的近縁性を反映している(値が大きいと縄文人に近い)可能性がある。

各都道府県とCHB間の f_2 統計量を計算したところ、47都道府県の中では奈良県が遺伝的にCHBに最も近く(f_2 値が最小)近畿地方と四国地方の県は、他の地方の都道府県よりもCHBに近かった。また、 f_2 統計量と図5のPC1は有意に相関していた(ピアソンの相関係数 $R=-0.88$ 、 P 値= 3.79×10^{-16})。現在のCHBが渡来人の遺伝的特徴を維持している(遺伝的に近縁)と仮定すれば、PC1は、縄文人と渡来人の2つの祖先集団との遺伝的類似性を反映している(PC1の値が大きければ縄文人に近く、小さければ渡来人に近い)と言えるだろう。

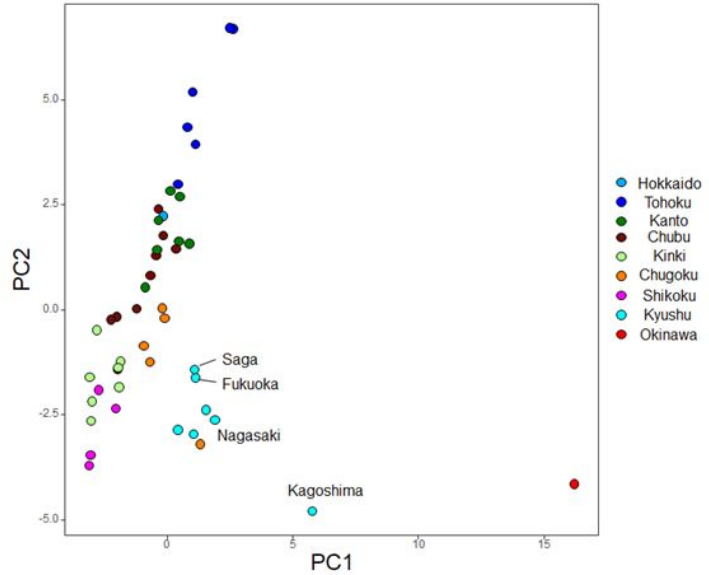


図5 47都道府県のPCAプロット

PCAの結果を可視的に理解するため、沖縄県を除いた46都道府県のPC1スコアの値に応じて色を付けた図を示す(図6)。図6において、青色が濃い県は遺伝的に縄文人に近く、オレンジ色が濃い県は渡来人に近い。東北地方と九州地方の県は縄文人の遺伝的要素を比較的好く残しており、近畿地方と四国地方の県は渡来人の遺伝的要素を残していると考えられる。

渡来人が導入した水田稲作が北九州で始まったことから、渡来人は最初に九州北部に到着した可能性が高い。しかし、九州地方北部(福岡・佐賀・長崎)の県に縄文人の遺伝的要素が比較的好く残っていたのは意外であった。渡来人が四国地方に行くには、九州地方もしくは中国地方を経由し、瀬戸内海を渡る必要がある。図6のオレンジ色が濃い地域と、前述の想定した渡来人の移動ルートは一致していないのは明らかである。オレンジ色の濃い地域で、渡来人の増加率が高かったのかもしれない。今後は、日本列島内で、渡来人と縄文人の混血がどのように進んだのかを調べていきたい。

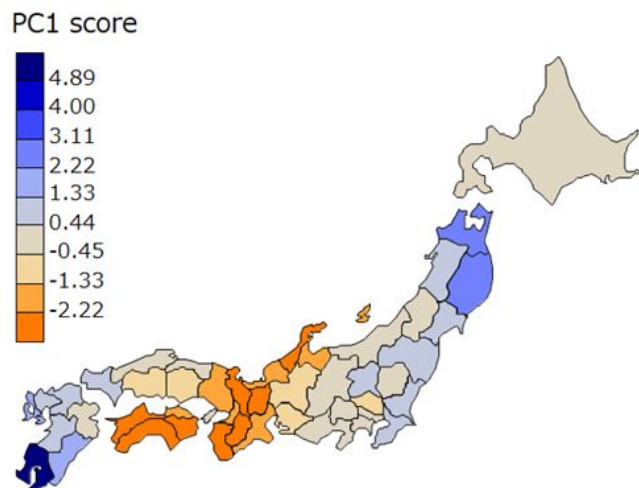


図6 PC1の値で色分けした日本地図(沖縄県除く)

5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 3件/うち国際共著 1件/うちオープンアクセス 0件）

| | |
|---|-------------------------|
| 1. 著者名 Watanabe Yusuke, Naka Izumi, Khor Seik-Soon, Sawai Hiromi, Hitomi Yuki, Tokunaga Katsushi, Ohashi Jun | 4. 巻 9 |
| 2. 論文標題 Analysis of whole Y-chromosome sequences reveals the Japanese population history in the Jomon period | 5. 発行年 2019年 |
| 3. 雑誌名 Scientific Reports | 6. 最初と最後の頁 8556 |
| 掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s41598-019-44473-z | 査読の有無 有 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |
| 1. 著者名 Cameron No?l, Yamauchi Taro, Ohashi Jun | 4. 巻 46 |
| 2. 論文標題 Human Biology of Japan | 5. 発行年 2019年 |
| 3. 雑誌名 Annals of Human Biology | 6. 最初と最後の頁 285 ~ 286 |
| 掲載論文のDOI（デジタルオブジェクト識別子） 10.1080/03014460.2019.1671675 | 査読の有無 無 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 該当する |
| 1. 著者名 大橋 順、渡部裕介 | 4. 巻 73 |
| 2. 論文標題 現代人のゲノムデータから過去の人口変動を推定する：縄文人が経験した急激な人口減少 | 5. 発行年 2019年 |
| 3. 雑誌名 遺伝：生物の科学 | 6. 最初と最後の頁 424 ~ 428 |
| 掲載論文のDOI（デジタルオブジェクト識別子） なし | 査読の有無 無 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |
| 1. 著者名 Watanabe Yusuke, Isshiki Mariko, Ohashi Jun | 4. 巻 66 |
| 2. 論文標題 Prefecture-level population structure of the Japanese based on SNP genotypes of 11,069 individuals | 5. 発行年 2020年 |
| 3. 雑誌名 Journal of Human Genetics | 6. 最初と最後の頁 431 ~ 437 |
| 掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s10038-020-00847-0 | 査読の有無 有 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |

| | |
|---|---------------------------|
| 1. 著者名 Isshiki Mariko, Watanabe Yusuke, Ohashi Jun | 4. 巻 140 |
| 2. 論文標題 Geographic variation in the polygenic score of height in Japan | 5. 発行年 2021年 |
| 3. 雑誌名 Human Genetics | 6. 最初と最後の頁 1097 ~ 1108 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s00439-021-02281-4 | 査読の有無 有 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |

〔学会発表〕 計5件 (うち招待講演 0件 / うち国際学会 0件)

| |
|---|
| 1. 発表者名 渡部裕介、覚張隆史、中伊津美、Khor Seik Soon、澤井裕美、人見祐樹、西田奈央、徳永勝士、太田博樹、大橋順 |
| 2. 発表標題 古代縄文人ゲノムおよび現代日本人集団ゲノムに基づく縄文人のHLA型の推定 |
| 3. 学会等名 第28回日本組織適合性学会大会 |
| 4. 発表年 2019年 |

| |
|------------------------------------|
| 1. 発表者名 渡部 裕介、大橋 順 |
| 2. 発表標題 現代日本人ゲノム中の縄文人由来候補SNPの検出 |
| 3. 学会等名 第73回日本人類学会大会 |
| 4. 発表年 2019年 |

| |
|-----------------------------|
| 1. 発表者名 久我 明穂、太田 博樹、大橋 順 |
| 2. 発表標題 主成分分析を用いた混血率の推定 |
| 3. 学会等名 第73回日本人類学会大会 |
| 4. 発表年 2019年 |

| |
|--|
| 1. 発表者名 大橋順, 渡部裕介, 中伊津美 |
| 2. 発表標題 理論集団遺伝学的視点からみた縄文人と渡来人の混血が日本人ゲノムに与えた影響 |
| 3. 学会等名 第66回日本人類遺伝学会大会 |
| 4. 発表年 2021年 |

| |
|--|
| 1. 発表者名 渡部裕介, 河合洋介, 大前陽輔, 宮原麗子, Seik-Soon Khor, 野入英世, 大橋順, 徳永勝士 |
| 2. 発表標題 現代日本人の集団ゲノムデータに基づく縄文人の全ゲノムSNPのアリル頻度の推定 |
| 3. 学会等名 第75回日本人類学会大会 |
| 4. 発表年 2021年 |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

| 氏名 (ローマ字氏名) (研究者番号) | 所属研究機関・部局・職 (機関番号) | 備考 |
|---------------------------|-----------------------|----|
|---------------------------|-----------------------|----|

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| 共同研究相手国 | 相手方研究機関 |
|---------|---------|
|---------|---------|