

令和 4 年 6 月 1 日現在

機関番号：32644

研究種目：基盤研究(B) (一般)

研究期間：2019～2021

課題番号：19H04211

研究課題名(和文) 機械学習を駆使したゲノム多型ビッグデータからのヒト顔形状の予測

研究課題名(英文) Prediction of human facial morphology based on genome polymorphisms

研究代表者

今西 規 (Imanishi, Tadashi)

東海大学・医学部・教授

研究者番号：80270461

交付決定額(研究期間全体)：(直接経費) 13,500,000円

研究成果の概要(和文)：ヒトの顔形状は遺伝率の高い形質であるため、原理的にはゲノム多型の情報を用いて形状を予測することが可能である。本研究チームはゲノム多型データからヒトの顔形状を予測するためのソフトウェア「ゲノム・モンタージュ」の開発に取り組んできたが、日本人の顔形状の3D測定データに基づく主成分得点を多遺伝子スコア(genome-wide polygenic scores)によって予測することにより、ゲノム・モンタージュの基本ソフトウェアを開発することに成功した。今後は顔形状およびゲノム多型データを数千人程度にまで増やし、さらにソフトウェアの改良を加えることで、顔形状の予測精度を高めることをめざす。

研究成果の学術的意義や社会的意義

学術的意義：生物の複雑な形態とゲノムとの関係は、古くから生物学における未解明の問題であった。本研究では、形態とゲノムの関係を解明するための有効な方法論を開発できたと考えられるので、大きな意義がある。また、本研究で用いた方法論は他の形態にも応用可能であり、発展性がある。

社会的意義：ヒトの顔形状予測ソフトウェアは法医学をはじめ考古学や人類学などの分野での応用が期待される。

研究成果の概要(英文)：Since the human facial morphology is a trait with a very high heritability, it is in principle possible to predict the facial morphology using information on genomic polymorphisms. Our research team has been working on the development of software "Genome Montage" for predicting human facial morphology based on genomic polymorphisms. Using 3-dimensional facial morphological data of the Japanese subjects, we succeeded in developing basic software of Genome Montage, by predicting principal component scores of the facial morphology from the genomic polymorphism data using genome-wide polygenic scores. In the future, we aim to improve the accuracy of predicted faces by increasing the number of subjects to about several thousands and by improving the software.

研究分野：生命情報学

キーワード：ゲノムモンタージュ 人類学 顔形状 機械学習 GWAS

## 1. 研究開始当初の背景

一卵性双生児はよく似た顔をしていることから、ヒトの顔形状の決定には遺伝要因が強く働いていると考えられる。そのため、多くの被験者に対する顔形状とゲノム情報を集めて解析を行うことができれば、ゲノム情報から顔形状の予測を行う情報技術(ゲノム・モンタージュ)の開発が原理的には可能である。しかし、その実現には多くの課題がある。第一に、複雑な顔面の形状を数値化して分析可能な指標にする必要がある。さらに、形状に関する複数の指標に対して複数遺伝子座との関連を調べるという多重検定の問題や、遺伝子座間だけでなく形状指標間の相互作用の問題に対処しなければならない。この意味で、顔の形状予測は生命情報学におけるチャレンジングな問題である。こうした研究は海外でも行われており、一定の成果を挙げているが、顔面形態を規定する遺伝要因の全容解明には至っていない状況にある。

代表者らは 2016-2018 年度に科研費・基盤 A の予算を受けて、研究課題「顔形状を規定するゲノム変異の網羅的探索によるゲノム・モンタージュ技術の開発」を実施してきた。この研究では、第一に日本人双生児 180 名の顔形状計測を行い、顔のさまざまな部位の遺伝率推定を行った。第二に、東北大学加齢医学研究所の収集した日本人ボランティアの頭部 MRI データを使い、顔形状を取り出して、ゲノムワイド関連解析を行った。これらの研究では画像データの解析を行うための各種ソフトウェアを整備してきた。しかし、これまでの海外の研究事例などと比較しても、集めたサンプル数が少ないことが大きな問題であった。そこで本研究では、多数の日本人の顔形状データとゲノム多型データを集めることに特に重点を置いた研究計画を立案した。

## 2. 研究の目的

本研究は、ゲノム情報に基づいてヒトの顔形状を予測するためのソフトウェア「ゲノム・モンタージュ」を作成することを主目的とする。ゲノム・モンタージュの実現のために、多くの日本人ボランティアからゲノム情報と顔形状データを取得し、計算機による大規模な関連解析を実施する。また、急速に発展しつつある機械学習のソフトウェアを活用し、ゲノム多型に基づいて 2 次元および 3 次元の顔形状の予測を行い、機械学習の有効性を評価する。これと並行して、東北大学加齢医学研究所が収集した頭部 MRI データとゲノム情報を解析し、顔形状に関わるゲノム多型の特定をめざす。

## 3. 研究の方法

データ収集のために新たに日本人成人ボランティアを募集し、3D スキャナによる顔表面形状の測定と口腔内粘膜 DNA の採取を行った。DNA は SNP アレイによるゲノム多型解析にかけ、ゲノムモンタージュ作成のためのデータを準備した。ボランティアには対面により研究内容を説明し、書面で同意を得た。取得した顔表面形状データ(点群データ)はノイズ除去などの前処理を行い、あらかじめ多数の解剖学的特徴点を定義したモデル顔(30 人の平均顔)との重ね合わせを行い、特徴点の位置(3 次元座標)を決定した。こうして、顔形状を個人間で比較したり、モデル顔との違いを数値化したりできるようになった。

3D データの解析およびゲノムモンタージュのためのソフトウェア開発は、主に統計解析ソフト R のスクリプトとして作成し、一連の解析プログラムをパッケージとして完成させることをめざした。各プログラムのマニュアルも整備した。

## 4. 研究成果

新たなデータの収集は、COVID-19 の流行のため規模を縮小して行うことになった。そのため、本研究では主にこれまでに収集済みのデータを使用して、データ解析とゲノム・モンタージュのソフトウェア開発に注力することとした。

3D データの解析においては、顔の表面形状の 3D 測定データを点の集合(点群)として扱った。モデル顔との重ね合わせ後に、すべての点の 3 次元座標の値を主成分分析にかけて第 1 から第 10 主成分を抽出することによって、顔の特徴をあらわす 10 個のパラメータを得た。これを「顔特徴パラメータ」と定義した。10 個の顔特徴パラメータにより、顔形状の分散の約 7 割を説明できることがわかった。次に、約 30 万ヶ所の SNPs を使い、被験者のゲノム多型データから多遺伝子スコア(genome-wide polygenic scores)を計算し、顔特徴パラメータの値を予測した。また、任意の値の顔特徴パラメータに対応する顔形状を 3D 表示するプログラムも用意した。以上

により、ゲノムモンタージュに必要な基本プログラムの開発に成功した。

多遺伝子スコアを用いた顔特徴パラメータの予測値は実際の値にかなり近く、形状の再現性が良いことが判明した。現時点ではデータ数が十分でないため統計学的な評価が十分にできていないが、顔全体の予測を行った場合には少なくとも顔の輪郭をよく再現できていることを確認できた。

また、顔全体を対象とした予測では顔の細部の形態の特徴を十分に再現できていない恐れがあるため、顔の点群データを目、鼻、口、その他の4つの部分に分けて（図1）、それぞれをゲノム多型データから形状予測するツールを開発した。これにより、部分ごとに形状を正確に予測し、その後に顔全体を組み立てる方法を実現できた。例えば鼻の形状では、実際の形状にかなり近い予測形状が得られている（図2）。

このほか、2D写真からの目の形状の分析を行った。3Dスキャナによる形状測定時には強い光を当てるため、閉眼状態で撮影しており、開眼状態での目の形は不明である。そこで、被験者の顔を正面から見た2D写真を用いて、開眼状態での目の形を分析した。写真上で目の形状の特徴をとらえるために約20ヶ所の特徴点を決め、その座標をもとに上瞼と下瞼の曲線および虹彩の位置を決めた。次に、上下の瞼で囲まれた領域の面積や目の縦横比、上瞼と下瞼の曲線の傾きなどのパラメータを使い、目の形の特徴をとらえる主成分を抽出し、個人ごとの目の形の特徴を数値化した（目の特徴パラメータ）。その結果、第1主成分は蒙古ひだを含む目の輪郭の特徴、第2主成分は目の縦横比（パッチリ目か細目か）に相当する成分が得られた（図3）。次は、これらの目の特徴パラメータをゲノム多型データから推定し、開眼状態での目の形を予測する予定である。また、2Dでの目の形状予測結果を3Dデータに組み込む方法の開発も計画している。

今後は顔形状およびゲノム多型データを数千人程度にまで増やし、さらにソフトウェアの改良を加えることで、顔形状の予測精度を高めることをめざす。なお、本研究期間中にCOVID-19が流行し、日本政府から緊急事態宣言が発令される状況となった。顔形状データとDNAサンプルを収集する際には被験者との接触が避けられないため、大規模な測定イベントの開催は実質的に不可能となり、当初予定していた数の顔形状データとDNAサンプルを集めることはできなかった。2022年の後半からは徐々に制約が緩和されると期待されるので、安全に配慮した方式での測定イベントを再開し、データ・サンプルの収集を実施していく方針である。

図1.部分ごとに分割した点群データ

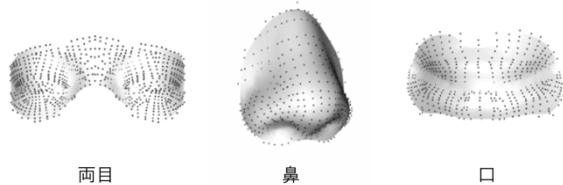


図2.鼻の形状の予測例

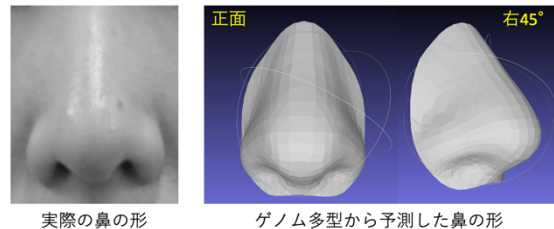
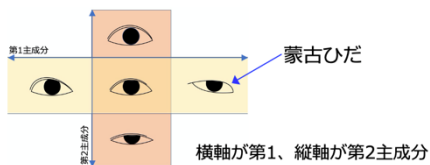


図3.目の形状の特徴パラメータ



## 5. 主な発表論文等

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

1. 著者名 Kataoka Keiichi, Fujita Hironori, Isa Mutsumi, Gotoh Shimpei, Arasaki Akira, Ishida Hajime, Kimura Ryosuke	4. 巻 11
2. 論文標題 The human EDAR 370V/A polymorphism affects tooth root morphology potentially through the modification of a reaction?diffusion system	5. 発行年 2021年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 5143
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s41598-021-84653-4	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 木村 亮介	4. 巻 24
2. 論文標題 東アジアにおけるヒトの遺伝適応の痕跡を求めて	5. 発行年 2019年
3. 雑誌名 日本生理人類学会誌	6. 最初と最後の頁 87～91
掲載論文のDOI（デジタルオブジェクト識別子） 10.20718/jjpa.24.2_87	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Takeuchi H, Kimura R, Tomita H, Taki Y, Kikuchi Y, Ono C, Yu Z, Matsudaira I, Nouchi R, Yokoyama R, Kotozaki Y, Nakagawa S, Hanawa S, Iizuka K, Sekiguchi A, Araki T, Miyauchi CM, Ikeda S, Sakaki K, Dos S Kawata KH, Nozawa T, Yokota S, Magistro D, Imanishi T, Kawashima R.	4. 巻 42
2. 論文標題 Polygenic risk score for bipolar disorder associates with divergent thinking and brain structures in the prefrontal cortex	5. 発行年 2021年
3. 雑誌名 Human Brain Mapping	6. 最初と最後の頁 6028～6037
掲載論文のDOI（デジタルオブジェクト識別子） 10.1002/hbm.25667	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Takeuchi Hikaru, Tomita Hiroaki, Taki Yasuyuki, Kikuchi Yoshie, Ono(Tanaka) Chiaki, Yu Zhiqian, Matsudaira Izumi, Nouchi Rui, Imanishi Tadashi, Kawashima Ryuta	4. 巻 in press
2. 論文標題 A psychiatric disorder risk polymorphism of <sc>ITIH3</sc> is associated with multiple neuroimaging phenotypes in young healthy adults	5. 発行年 2022年
3. 雑誌名 Psychiatry and Clinical Neurosciences	6. 最初と最後の頁 in press
掲載論文のDOI（デジタルオブジェクト識別子） 10.1111/pcn.13347	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Yamaguchi Ryo, Matsudaira Izumi, Takeuchi Hikaru, Imanishi Tadashi, Kimura Ryosuke, Tomita Hiroaki, Kawashima Ryuta, Taki Yasuyuki	4. 巻 494
2. 論文標題 RELN rs7341475 Associates with Brain Structure in Japanese Healthy Females	5. 発行年 2022年
3. 雑誌名 Neuroscience	6. 最初と最後の頁 38 ~ 50
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neuroscience.2022.05.007	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計7件 (うち招待講演 3件 / うち国際学会 2件)

1. 発表者名 Hikaru Takeuchi
2. 発表標題 Polygenic risk score for bipolar disorder associates with divergent thinking and brain structures in the prefrontal cortex
3. 学会等名 Society for neuroscience of creativity, virtual meeting (国際学会)
4. 発表年 2020年

1. 発表者名 木村亮介
2. 発表標題 東ユーラシア系ヒト集団にみられる表現型多様性と適応
3. 学会等名 国立遺伝学研究所研究会「日本列島人の起源と成立をゲノム情報から探る」(招待講演)
4. 発表年 2021年

1. 発表者名 Kimura R
2. 発表標題 Pleiotropic effects of the EDAR 370V/A variant on East Asian phenotypes.
3. 学会等名 Genome Concept Centennial Conference, web meeting (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 木村亮介、宮平由香子、金城優奈、伊佐睦実、小金淵佳江、石田肇、安藤寿康、中川草、今西規
2. 発表標題 幾何学的形態測定学を用いた三次元顔面形態の解析と遺伝率推定
3. 学会等名 第73回日本人類学会大会（招待講演）
4. 発表年 2019年

1. 発表者名 今西規、中川草、木村亮介、瀧靖之、竹内光、安藤寿康
2. 発表標題 個人ゲノム情報に基づくヒト顔形状の予測
3. 学会等名 第42回日本分子生物学会年会
4. 発表年 2019年

1. 発表者名 今西 規、中川 草、木村 亮介、瀧 靖之、竹内 光、安藤 寿康
2. 発表標題 個人ゲノム情報に基づくヒト顔形状の予測
3. 学会等名 第75回日本人類学会大会
4. 発表年 2021年

1. 発表者名 後藤新平、片岡恵一、石田肇、木村亮介
2. 発表標題 骨の厚さに関連する要因：頭蓋骨と上腕骨の比較
3. 学会等名 第127回日本解剖学会総会・全国学術集会 2022年
4. 発表年 2022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
研究 分担者	木村 亮介  (Kimura Ryosuke)  (00453712)	琉球大学・医学(系)研究科(研究院)・准教授   (18001)	
研究 分担者	瀧 靖之  (Taki Yasuyuki)  (10375115)	東北大学・スマート・エイジング学際重点研究センター・教授   (11301)	
研究 分担者	竹内 光  (Takeuchi Hikaru)  (50598399)	東北大学・加齢医学研究所・准教授   (11301)	

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
---------	---------