

## 論文内容の要旨

### 論文題目

A scan for genetic determinants of human hair morphology using genome-wide SNP database

### 和訳

ゲノムワイド SNP データベースを用いた毛髪の形態決定遺伝子の探索

指導教官 徳永勝士教授

東京大学医学系研究科

平成17年4月入学

博士後期課程

国際保健学専攻

氏名 藤本明洋

### 緒言

毛髪の形態は、人類集団間で最も大きく分化している形質の一つである。アフリカ系集団には細い縮毛の個体が多く、ヨーロッパ系の集団には細い波状毛の個体が多く、アジア系の集団には太い直毛の個体が多いことが知られている。また、遺伝学的研究によって、毛髪の形態は少数の遺伝子によって決定されていると推測されている。しかしながら、毛髪の形態に関与する遺伝子は未だに同定されていない。さらに、人類集団間の毛髪の形態的差異をもたらした進化的要因も解明されていない。

本研究において、私は、(1) 最初に、毛髪の形成に関与する170の遺伝子を対象に、国際ハップマップ計画により公開されたゲノムワイド SNP データベースを用いて候補遺伝子の絞り込みを行った。(2) 次に、絞り込まれた候補遺伝子内の SNP と毛髪の形態との関連を検討した。さらに、(3) 関連が認められた SNP について機能解析を行なうとともに、(4) 集団遺伝学的アプローチを用いて、進化学的解析を行なった。その結果、*EDAR* (*ectodysplasin A receptor*) がアジア人の毛髪の太さに関与し、最近の自然選択によって進化した事が示唆された。

## 対象と方法

### 集団間の遺伝的分化による候補遺伝子の選出

国際ハップマップ計画によって公開された SNP データ（ハップマップデータベース）を解析に用いた。ハップマップデータベースでは、3 集団に由来する 2 1 0 人（アフリカ系 6 0 人、ヨーロッパ系 6 0 人、北東アジア系 9 0 人）のゲノムワイドな SNP データが公開されている。

最初に、人またはマウスにおいて毛髪形成に關与する事が知られている遺伝子（毛髪関連遺伝子）を 1 7 0 個選出した。次に、それらの遺伝子の配列を 5 0 kb の領域に分割し、それぞれの領域における  $F_{ST}$  の最大値 ( $mF_{ST}$ ) を求めた ( $F_{ST}$  は集団間の遺伝的分化の程度を表す指標であり、遺伝的分化が完全である場合は 1 となり、全く分化していない場合は 0 となる)。毛髪の形態は人類集団間で大きく分化し、少数の遺伝子によって決定されていることが示唆されているため、毛髪の形態決定遺伝子は特に大きな集団間分化を示すと期待される。その後、全ゲノムについても  $mF_{ST}$  を計算し、ゲノムワイドな  $mF_{ST}$  の経験分布を求めた。経験分布のもとで上位 1 % または 5 % の高い集団間の遺伝的分化を示す毛髪関連遺伝子を候補として選出した。

### 対象

毛髪の形態と遺伝子の関連を調査するため、タイとインドネシアにおいて、毛髪および DNA サンプルを収集した。タイにおいては、モーケン族およびウラクラオイ族からなる集団の 6 5 人を対象とした。インドネシアにおいては、スンダ族およびジャワ族からなる集団の 1 2 1 人を対象とした。

毛髪サンプルは後頭部より採取し、一人当たり 5 本の毛髪の垂直切片を作成し、顕微鏡下で長径、短径および面積を測定した。また、長径と短径の比 (hair index) を計算し、毛髪断面の形状の評価に用いた。EDAR 上の 1540T/C および -1430G/A 多型の遺伝子型を PCR ダイレクトシーケンス法によって決定し、毛髪の形態との統計学的関連を解析した。

さらに、1540T/C の遺伝子頻度の地理分布を推定するため、タイ、インドネシア集団に加え、オセアニアの 2 集団（パプアニューギニアのギデラ族、ソロモンの地域集団）それぞれ 4 8 個体の遺伝子型を決定した。

### ルシフェラーゼアッセイ

EDAR は下流の転写因子 NF- $\kappa$ B を活性化することが知られている。EDAR 1540T/C のア

レル間の機能的差異を検証するため、HeLa 細胞および 293A 細胞を用いてルシフェラーゼアッセイを行なった。両アレルの発現ベクターを作製し、NF- $\kappa$ B が結合するエンハンサーエレメントとルシフェラーゼ遺伝子を持つベクターともに細胞に導入し、48 時間後にルシフェラーゼ活性を測定した。

## 進化学的解析

*EDAR* 1540T/C における自然選択の有無を検証するため、LRH テスト (long range haplotype test) を行なった。LRH テストは、短い期間で頻度を増したハプロタイプは歴史的組換え数 (誕生してから組換えを受けた回数) が少ないことを利用し、最近の自然選択の証拠を提示する方法である。LRH テストでは、それぞれのアレルについてハプロタイプの同一性 (EHH: Extended Haplotype Homozygosity) を計算し、対象とするアレルと他方のアレルの EHH の比 (REHH: relative EHH) を用いて検定を行なう。対象アレルのハプロタイプ同一性が一般的なレベルよりも高ければ (REHH の値が有意に高ければ)、最近の自然選択の影響によって急速に頻度を上昇させたと考えられる。私は、C アレルについての REHH を、*EDAR* が位置する 2 番染色体の REHH の経験分布を用いて検定した。LRH テストには、ハップマップデータベースにより公開された北東アジア集団 (CHB+JPT) のハプロタイプデータを用いた。

## 結果と考察

### 毛髪形成に關与する遺伝子の集団間分化

ゲノムワイドな  $mF_{ST}$  の経験分布と、それらの 95 パーセンタイルおよび 99 パーセンタイルを図 1 に示す。170 個の毛髪関連遺伝子の中で、21 遺伝子が上位 5% 以上の遺伝的分化を、5 遺伝子が上位 1% 以上の遺伝的分化を示した。上位 1% 以上の遺伝的分化を示した遺伝子のうち、*EDAR* (*ectodysplasin A receptor*) は北東アジア集団 (CHB+JPT) とその他の集団 (ヨーロッパ集団 (CEU) とアフリカ集団 (YRI)) の間で高い遺伝的分化を示した。また、遺伝子上に非常に高い遺伝的分化を示す非同義多型 (1540T/C) が観察された。この多型において、C アレルはヨーロッパ集団およびアフリカ集団では観察されなかったが、北東アジア集団では 80% 以上の頻度に達し

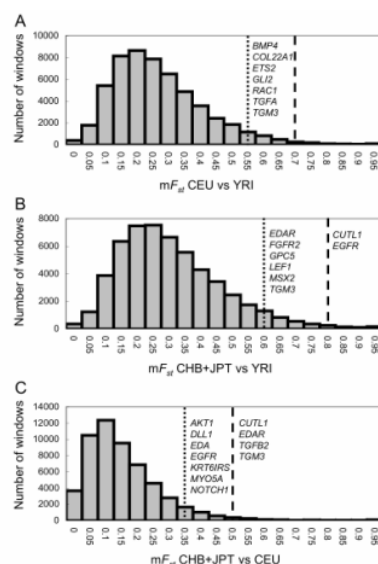


図1  $mF_{ST}$  の経験分布と高い集団間分化を示した毛髪関連遺伝子  
(A) ヨーロッパ集団 (CEU) vs. アフリカ集団 (YRI)  
(B) 北東アジア集団 (CHB+JPT) vs. アフリカ集団 (YRI)  
(C) 北東アジア集団 (CHB+JPT) vs. ヨーロッパ集団 (CEU)  
点線は95パーセンタイル、棒線は99パーセンタイルを示す。

ていた。以上より、*EDAR*を候補遺伝子とし、以下の解析を行なった。

### *EDAR*と毛髪の形態の関連解析

タイおよびインドネシアで採取したサンプルの、*EDAR*上の非同義多型(1540T/C)およびプロモーター領域の多型(-1430G/A)の遺伝子型を決定し、毛髪の形態と

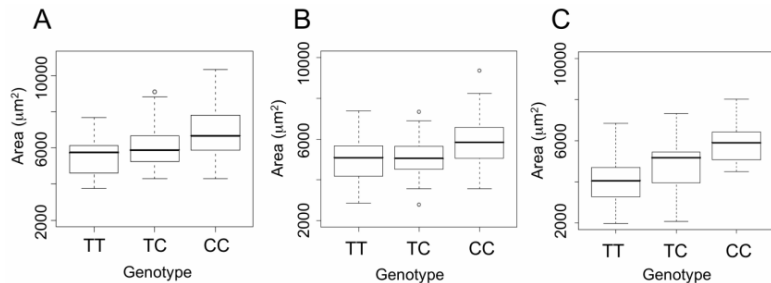


図2 *EDAR* 1540T/Cと毛髪の断面積の関連 (A) 日本、(B) インドネシア、(C) タイ

の関連解析を行なった。-1430G/Aは毛髪の形態との有意な関連を示さなかったが、1540T/Cは、長径、短径および面積と有意に関連していた。特に、面積は強い関連(ANOVA 日本  $P=1.4 \times 10^{-5}$ 、インドネシア  $P=5.5 \times 10^{-3}$ 、タイ  $P=9.5 \times 10^{-4}$ )を示した(図2)。また、アジア特異的なCアレルは面積を増加させる効果があることも示唆された(図2)。さらに、年齢、民族差、性別について重回帰分析を行なった結果、1540T/Cに加えて年齢も毛髪の面積に関連する事が分かった。

以上の結果より、*EDAR*はアジア人の毛髪の太さの決定遺伝子であると考えられる。

### 機能解析

*EDAR*は転写因子NF- $\kappa$ Bを活性化することにより、下流の遺伝子発現を調節し、毛包の形成に重要な役割を果たしていることが知られている。したがって、*EDAR* 1540T/Cの機能的差異がNF- $\kappa$ Bの活性化に影響する可能性が考えられる。この仮説を検証するため、私はルシフェラーゼアッセイを行なった。

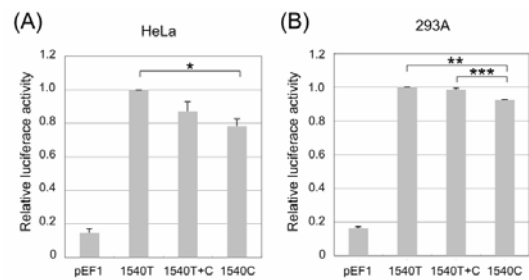


図3 ルシフェラーゼアッセイ (A) HeLa細胞、(B) 293A細胞

その結果、Cアレル導入細胞では、Tアレル導入細胞に比べ、ルシフェラーゼ活性の有意な低下が観察された(図3)。この結果は、1540T/Cは下流のNF- $\kappa$ Bの活性に影響を与え、毛包内の細胞の分化または増殖に影響を与えることを示唆する。

### 進化学的解析

*EDAR* 1540T/Cの進化史を推定するため、1540T/Cの遺伝子頻度の地理分布を求め、LRHテストを行なった。Cアレルは、特に北東アジアにおいて高い頻度を示した(図4A)。

また、オセアニア集団ではCアレルが低い頻度で存在した（図4A）。さらに、LRHテストの結果、CアレルのREHHの値は2番染色体のREHHの分布と比較して、有意に高いことが示された（図4B）。これらの結果より、Cアレルはアジア集団がヨーロッパ集団から分岐した後に生じ、北東アジアにおいて地域特異的な自然選択の影響を受けて急速に頻度を上昇させたことが示唆される。

Cアレルにおける自然選択の原因は不明であるが、（1）Cアレルは太い毛髪に関連することから、北東アジア集団においては太い毛髪をもつ個体の方が寒さから身を守る上で有利であった可能性や、（2）EDARは毛髪に加えて歯や汗腺の形成にも関与しているため、自然選択は毛髪以外の形質（例えば歯の形態など）に作用し、二次的に毛髪の形態の変化を引き起こした可能性などが考えられる。

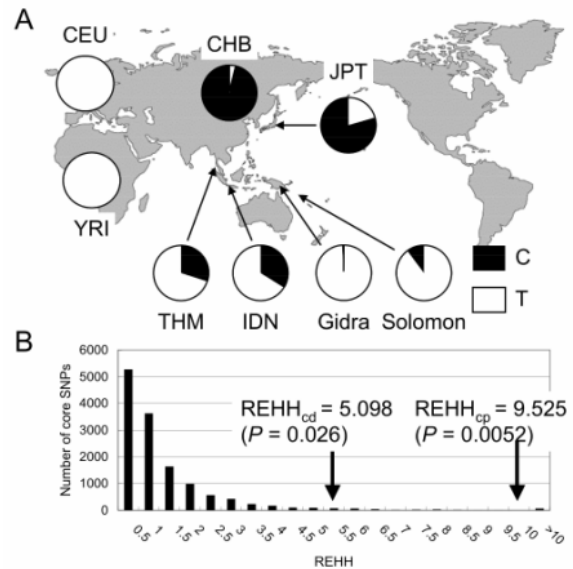


図4 EDAR 1540T/Cにおける自然選択  
 (A) 1540T/Cの地理分布  
 東南アジア (THM;タイ, IDN;インドネシア)、オセアニア (Gidra, Solomon)、北東アジア (JPT, CHB)、ヨーロッパ (CEU)、アフリカ (YRI)  
 (B) LRHテスト  
 REHH<sub>cd</sub>: 下流0.25cMにおけるREHH  
 REHH<sub>cp</sub>: 上流0.25cMにおけるREHH