

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 22 日現在

機関番号：82111

研究種目：基盤研究(B) (一般)

研究期間：2013～2016

課題番号：25292190

研究課題名(和文) カスタム育種デザインのためのブタ遺伝子発現調節領域の多様性の解明

研究課題名(英文) Analysis of polymorphisms in regulatory regions of gene expression for pig custom breeding design

研究代表者

上西 博英 (Uenishi, Hirohide)

国立研究開発法人農業・食品産業技術総合研究機構・生物機能利用研究部門・動物機能利用研究領域・ユニット長

研究者番号：80391556

交付決定額(研究期間全体)：(直接経費) 14,900,000円

研究成果の概要(和文)：豚の脂質に関する改良の指標策定のために、脂質合成や、脂肪細胞の分化・増殖に関連する遺伝子の発現を調節しているプロモータ領域の中の塩基配列の違い(多型)を、ゲノム濃縮と次世代シーケンサーを用いた解析により明らかにした。また、それらの多型の中で、プロモータに結合し遺伝子発現を調節する転写因子の内、脂質や脂肪細胞に関連するものが結合するゲノム上の領域で、転写因子の結合に影響することが想定されるものの抽出を行った。これらの多型は豚の脂肪の厚さ等の形質を支配するゲノム領域中に存在しているものもあり、脂肪に関する形質を改良するための指標としての利用可能性が示された。

研究成果の概要(英文)：To determine the indices to improve fat-related traits in pigs, we clarified polymorphisms in promoter regions of the genes related to lipid synthesis and/or differentiation and proliferation of adipocytes by genome concentration and next-generation sequencing techniques. Furthermore, we extracted the polymorphisms affecting binding ability of transcription factors related to lipid synthesis or adipocytes. Some of the extracted polymorphisms were located within the genomic regions of quantitative trait loci of fat-related phenotypes, suggesting the availability of the polymorphisms for improvement of fat-related traits of pigs.

研究分野：ゲノム科学

キーワード：ゲノム プロモータ 多型 ブタ 次世代シーケンサー 転写因子

1. 研究開始当初の背景

家畜においては、ウシ及びニワトリについてゲノム塩基配列の概要解読の成果が公開されており、ブタについても、研究担当者らも主要な役割を果たした国際ブタゲノムシーケンシングコンソーシアムによりゲノム塩基配列の概要解読が完了している。本概要解読では、ブタの 18 対の常染色体と性染色体について、全体のおよそ 90% をカバーする領域のゲノム塩基配列を解読することに成功している。また、研究担当者らは完全長 cDNA ライブラリーを用いたブタ遺伝子配列の解明に取り組み、33 万個以上の cDNA クローンの 5' 末端配列 (EST) を取得するとともに、それらの cDNA クローンの内 31,079 個については挿入配列の全長解読を完了している。これらのブタ遺伝子配列解読の成果は、ブタの遺伝子配列や構造についての世界で最も優れたリソースであり、概要解読が行われたブタゲノム塩基配列上における遺伝子の構造の推定に非常に大きな役割を果たしている。

これらのブタ等の家畜動物のゲノム塩基配列や遺伝子構造の情報は、分子生物学的な基盤情報として非常に重要であり、DNA 情報を用いた品種改良の加速化に貢献すると考えられる。ところが、家畜はマウス等の実験動物とは異なり同一品種内においても遺伝的多様性が高度に残存している。これらの塩基配列の多様性は、実験動物としての利用においては個体間差による実験結果のバラツキとして、また食肉等の畜産物生産においても品質が一定しないという点で好ましくない影響を与えるものであるが、品種改良という点ではまだまだ多くの余地を残していると考えられる。

2. 研究の目的

ブタにおいては、他の動物種と同様、遺伝的多様性がゲノム塩基配列上に広く分布しているが、コード領域内に存在し、タンパク質を構成するアミノ酸を変化させるものについては、形質に直接的に何らかの影響を与えることが推定される。しかしながら、タンパク質の構造を変化させるような変異は生存上大きく不利であることが多いと考えられ、現在残存している多型の中で品種改良に用いることができるものは必ずしも多くない可能性が想定される。一方、コード領域以外に存在する多型、特にプロモータ等の遺伝子の発現調節に関わる領域内に存在し、遺伝子の発現量を通じて様々な形質に影響を与えているものは、直接的に個体生存の有利・不利に結びつかないものが多いことが想定され、品種改良のための素材として有望であると考えられる。実際に、家畜においてこれまでに解析された量的形質遺伝子座 (QTL) において、アミノ酸置換ではなく、遺伝子発現に関わる領域の多型がその本体であることが推定されるものがいくつか報告されて

いる。

ブタにおいても、ゲノム塩基配列解読と、遺伝子配列の網羅的収集の成果によって、プロモータ等の発現調節領域の配列を詳細に解析できる環境が整いつつある。特に、完全長 cDNA ライブラリーに由来する EST 及び cDNA 全長配列は、実際にゲノム上で転写が開始される場所からクロニングされている蓋然性が高く、ゲノム塩基配列上の転写調節領域を推定するための有力な情報である。また、これらの転写調節領域における多型の内、実際に遺伝子の発現量に対して影響を与えているものは、品種改良において、特定の方向に形質を変化させるための有力な指標となり得る。一つ一つの遺伝子の影響は小さくても、このような指標を組み合わせることで、任意の方向に改良された品種の造成、即ち「カスタム育種デザイン」が可能になることが想定される (図 1)。

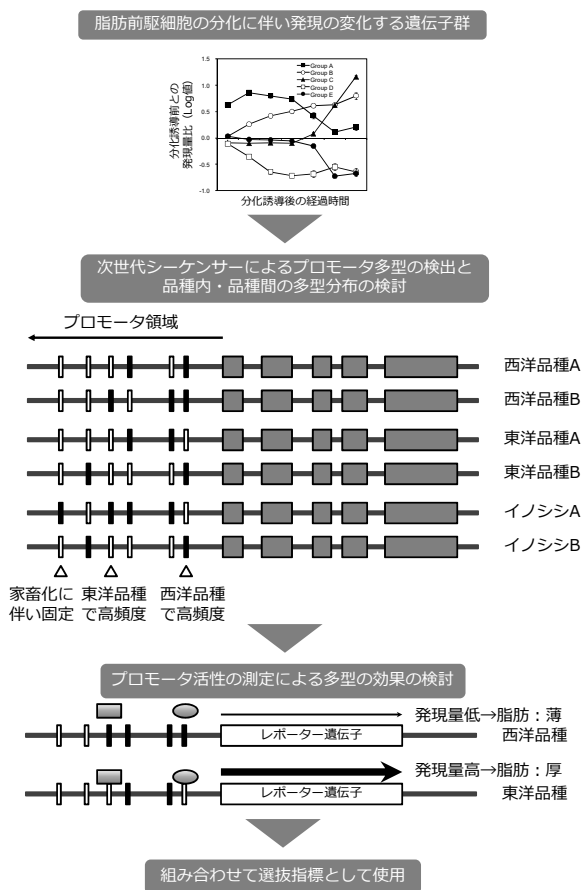


図 1. 脂質関連形質に関するカスタム育種デザインの概念図。

研究担当者らは、ブタの皮下由来の脂肪前駆細胞を用いて、脂肪細胞の分化に伴う遺伝子発現の経時的な変化の解析を行い、分化誘導刺激前と比較して 5 倍以上の発現量差を示す遺伝子が 300 個以上存在することを示している。2 倍以上の発現差を有意に示す遺伝子も、およそ 1,500 個程度検出している。5 倍以上の発現量差を示す遺伝子の内 20 個以上

については、既知の脂質関連形質の QTL 領域中に位置しており、これらの形質の責任遺伝子となっている可能性も示唆される。これらの脂肪細胞の分化と関連する遺伝子の発現調節領域は、遺伝的に皮下脂肪の厚さが異なる品種間で、特に転写因子が結合する領域において異なっていることが想定され、ブタの脂肪に関連する形質の改良において有力な指標となることが考えられた。

そこで、これまでに研究担当者が蓄積したブタ EST 情報及び cDNA クローン全長配列情報を最大限に活用し、ブタにおけるプロモータ領域の抽出と、主要商用豚品種等における転写調節に関わる領域中に含まれる遺伝的多様性を明らかにすることにより、ブタの品種改良の指標となる情報を提供することを主たる目的とする。モデル系としてブタ皮下由来脂肪前駆細胞の分化に伴い発現の変化する遺伝子を対象として解析を行う。これにより、脂質に関連する形質における「カスタム育種デザイン」に必要な情報が提供されることが期待される。

3. 研究の方法

ブタ皮下脂肪由来の脂肪前駆細胞に脂肪細胞への分化誘導を行った際に発現が変化する遺伝子群を対象として、転写開始点上流の配列を主要商用品種や東洋種のブタや、野生種（イノシシ）で比較することにより、皮下脂肪厚等、脂肪に関する形質とプロモータの多型との関連について推定を行うとともに、遺伝子発現に影響を与えていると考えられる多型については、脂肪に関して異なる形質を持つブタ集団間での遺伝子発現の比較等を行うことで、プロモータ多型の品種改良への利用における有効性の検討を行うこととした。

(1) ブタゲノム概要塩基配列からの脂肪細胞分化に関わる遺伝子のプロモータ領域の抽出

脂肪細胞の分化に関わる遺伝子群を対象に遺伝子発現に影響を与える可能性のあるプロモータ多型の検討を行うこととした。そのような遺伝子の候補として、遺伝子の機能として脂肪合成や代謝等との関連性があることが知られている遺伝子だけではなく、脂肪細胞の分化において、発現が上昇あるいは低下している遺伝子についても注目し、ブタにおいて重要な形質である脂肪の蓄積に関連していることが想定される遺伝子の抽出を行うこととした。抽出を行った遺伝子に対して、転写開始点の推定を行い、その上流配列をプロモータと見なして多型検索の対象候補とした。

転写開始点については、これまでに課題担当者が解読したブタ cDNA 全長解読の情報を中心に、GenBank 等の公的データベースに登録されている RefSeq 等の mRNA の情報を利用して同定を行う。完全長が解読されてい

ない遺伝子についても、完全長 cDNA ライブラリーが転写開始点からクローニングされている可能性が高いことを利用し、これまでに蓄積した EST 情報に基づいて転写開始点の推定を行った。

さらに、プロモータと想定され、転写調節に関わる配列を含む可能性が高いと想定されるゲノム上の領域（合計～10Mb）について、ゲノム DNA 濃縮のためのプローブセットの設計を行った。

(2) プロモータ領域中のブタ主要商用品種並びに野生種（イノシシ）中での次世代シーケンサーを用いた多様性検索

設計を行ったプローブセットを用い、国内で供用されているブタの主要商用西洋品種（デュロック・ランドレース・大ヨークシャー・バークシャー）、皮下脂肪が西洋種と比較して厚い東洋種（梅山豚、金華豚）について、次世代シーケンサーを用いたシーケンシングを行い、ブタゲノムリファレンス配列へのマッピングを行うことで、品種間、個体間の多型の抽出を行い、脂肪に関する形質の品種間での差異と、プロモータ多型との関連についての検討を行うこととした。さらに、プロモータ領域中で転写調節因子が結合すると想定される場所を、データベースを利用して抽出し、ブタ（イノシシ）で検出された、脂肪細胞分化に関わる遺伝子のプロモータ領域における多型の転写調節因子の結合性への影響についての検討を行った。

4. 研究成果

(1) ブタゲノム概要塩基配列からの脂肪細胞分化に関わる遺伝子のプロモータ領域の抽出

脂肪細胞分化あるいは脂肪合成等にかかわる遺伝子として、ヒトやマウス等における研究から脂肪合成等に関与することが示されている遺伝子群を抽出するとともに、ブタ脂肪前駆細胞の脂肪細胞への分化に伴い発現量が大きく変化する遺伝子に着目することとした。

既に研究分担者（中島）が樹立しているブタ皮下脂肪由来の脂肪前駆細胞（PSPA）に対して分化誘導を行い、誘導後 1 時間～6 日経過した細胞において 2 倍以上の発現変化を有意に示す遺伝子を、ブタ発現遺伝子配列に基づいて設計したオリゴマイクロアレイ（AGPOA2）を用いて検索を行ったところ、2,548 種（遺伝子としてはおよそ 1,500 個に相当）のマイクロアレイ上のプローブに対応する mRNA について発現変化が確認された。また、脂肪合成等に関与する遺伝子群として、Gene Ontology 中の脂質代謝（GO:0006629）、脂肪生合成（GO:0008610）、脂肪細胞分化（GO:0045444）、脂肪細胞増殖（GO:0070341）に一致するブタ遺伝子、あるいはヒト遺伝子で上述の Gene Ontology の項目に一致するもののブタでの相同遺伝子について等を抽出

することで、さらに2,508個の遺伝子（あるいはブタゲノム上の遺伝子と想定される領域）を抽出した。

これらの合計のべ5,056個のブタゲノム上の遺伝子（あるいは遺伝子相当領域）について、転写開始点を含むと想定される mRNA 全長配列を公的データベース上、あるいは研究代表者らが保有する完全長 cDNA ライブラリーに基づくブタ発現遺伝子データベース（Pig Expression Data Explorer; PEDE）を用い、mRNA 全長配列がないものについても mRNA 先頭からの端読みである EST データを活用して、さらに、国際コンソーシアムにより作成された、最新のブタゲノムリファレンス配列である Sscrofa10.2（基本的に西洋品種であるデュロック種個体のゲノム配列）と照合することにより転写開始点のゲノム上での位置の推定を行った。実際の転写開始点の推定に当たっては、① NCBI データベースで mRNA が登録されているもの（多くは研究代表者らの PEDE データベースに基づく）、② PEDE データベースに登録されているブタ EST 情報等、③ ブタで mRNA 全長及び EST 情報がない場合にヒトの mRNA をブタゲノムリファレンス配列（Sscrofa10.2）上にアラインメントの順で検討を行った。これらの方法によっても転写開始点の推定が困難であったものについては、AGPOA2 でプローブ設計をするに当たって利用した mRNA ないし EST 配列の先頭を転写開始点と見なした。これらにより、5,056 個のゲノム上の領域の内、重複を除いて3,041 箇所（12.9Mb に相当）について、10.2Mb のゲノム領域を濃縮可能な IonPGM™ 用のゲノム濃縮用プローブセット（Ion TargetSeq™による）を設計した。

(2) プロモータ領域中のブタ主要商用品種並びに野生種（イノシシ）中での次世代シーケンサーを用いた多様性検索

多型検索の対象として、国内で広く供用されている西洋品種 30 頭（パークシャー種 5 頭、デュロック種 8 頭、ランドレース種 9 頭、大ヨークシャー種 8 頭）、中国系（東洋）品種 11 頭（金華豚 4 頭、梅山豚 7 頭）、さらには東洋系のイノシシ（ニホンイノシシ 5 頭、リュウキュウイノシシ 3 頭）の合計 49 頭を選択した。これらのゲノム DNA を上述のゲノム濃縮用プローブセットを用いて検索対象領域について濃縮した上で、IonPGM™ システムを用いたシーケンシングを行った。得られたシーケンシングデータは BWA によりゲノムリファレンス配列 Sscrofa10.2 上にマッピングし、GATK により多型検出を行った。GATK での quality value が 100 以上、かつ同一種の塩基の連続が 5 個以下の部分に限定した信頼性の高いものに限定すると、多型部位として一塩基多型（SNP）は 109,869 箇所、挿入・欠失（indel）については 55,693 箇所が検出された。西洋品種中ではほぼ固定（対立遺伝子頻度が 5%以下）したもののうち、東洋品種に

おいて西洋品種で低頻度の遺伝型が 70%以上のもの（東洋型で特徴のある遺伝型を示す多型）は SNP で 4,571 箇所、indel で 1,562 箇所であった。さらに、西洋品種での高頻度遺伝型が東洋品種でも 5%以下となる、西洋-東洋品種間でほぼ分離しているものは SNP で 250 箇所、indel で 27 箇所存在した。

これらの多型、特に西洋-東洋品種で異なる多型については、東洋品種で顕著な脂肪の蓄積と遺伝子の発現調節を通じて関連性を有していることが推定される。そこで、今回多型検索を行った、脂質関連遺伝子のプロモータ領域と推定される配列中での、脂質合成・脂肪細胞分化等と関連性を有すると考えられる転写因子の結合配列の検索を行い、その中に含まれる多型の抽出を行うこととした。

転写因子の内、① 脂肪前駆細胞 PSPA の分化に伴い発現が 5 倍以上変化する遺伝子の発現に関わることが Gene Set Enrichment Analysis により示されたもの、② PSPA の分化において、転写因子をコードする遺伝子自身の発現が 5 倍以上変化したもの、③ KEGG データベース中で脂肪細胞と関連するパスウェイにおいて脂質合成と関連する転写因子を候補として抽出し、合計 14 個の転写因子結合配列を選択した。選択した転写因子結合配列については、ブタにおいては現在のところデータベースが存在しないため、転写因子結合配列が種間で高度に保存されていることを勘案し、転写因子結合配列データベース JASPAR2016 のヒトでのデータを活用してゲノム上の分布の検索を行った。

東洋品種と西洋品種での脂肪蓄積の違いが転写因子結合配列の違いに起因しているとするならば、転写因子結合配列内の多型により転写因子への結合能に変化が生じている可能性が考えられる。そこで、上記の「東洋型で特徴のある遺伝型を示す多型」について、UCSC ブタゲノム配列データベース（BSgenome.Sscrofa.UCSC.susScr3）を用いて、ブタゲノムリファレンス配列から東洋系で高頻度な遺伝型に変換した、「仮想東洋品種ゲノム配列」の構築を行った。さらに、ブタゲノムリファレンス配列と、仮想東洋品種ゲノム配列それぞれにおいて、TFBS tools を用いて R 環境上で上述の 14 種の転写因子結合配列の分布について検討した。その結果、ゲノム濃縮対象領域中で、ブタゲノムリファレンス配列に 90%以上の相同性で脂質関連転写因子結合配列として検出された箇所が SNP で 17,007 箇所・indel で 18,591 箇所、仮想東洋品種中では SNP で 17,019 箇所・18,623 箇所存在した。これらの中で、仮想東洋品種ゲノム配列にのみ転写因子の結合が想定される（ブタゲノムリファレンス配列で相同性を 80%としても結合が想定されない）ものは 9 箇所（すべて SNP）、ブタゲノムリファレンス配列でのみ結合が想定される（仮想東洋品種で相同性を 80%としても結合が想定されな

い) ものは2箇所 (SNPが1箇所・indelが1箇所) であった (表1)。

表1. 西洋品種由来のブタゲノムリファレンス塩基配列 (Sscrofa10.2) と仮想東洋品種ゲノム塩基配列との間で転写因子結合部位が異なる多型の一覧。今回ゲノム濃縮とシーケンシングを行った範囲内のみを示す。○を付してある方が該当する転写因子の結合が想定される配列である。

染色体	転写因子	遺伝子	Sscrofa10.2	仮想東洋 品種ゲノム
SSC2	NFATC2	不明	×	○
SSC2**	EGR1	不明	○	×
SSC6	CEBPD	不明	×	○
SSC7*	SREBF1	不明	×	○
SSC8	MEF2C	SGMS2	×	○
SSC9	EGR1	PIGC	○	×
SSC14	CEBPD	SIRT1	×	○
SSC18	CEBPD	不明	×	○
SSCX*	NFATC2	GDPD2	×	○

*2箇所の多型を同一ゲノム領域内に含む。

**本多型のみ indel である。他はすべて SNP である。

これらの中で、特に SSC2 (EGR1)、SSC7 (SREBF1)、SSC9 (EGR1) 及び SSC14 (CEBPD1) (括弧内は転写因子を示す) の多型については、転写開始点と想定される位置から 1kb 以内に該当の多型が存在し、下流の遺伝子の発現に対して強い影響を与えていることが想定される。これらの転写因子の結合に影響を与えることが想定される多型は、脂質合成、あるいは脂肪細胞の分化・増殖に影響を与えていることが想定され、脂肪に関する形質に着目した育種においてマーカーとなり得る可能性が高いと考えられる。例えば、SSC7 の該当多型については、脂肪の厚さや脂肪細胞の直径に関する数多くの量的形質遺伝子座 (QTL) が報告されている領域内に位置しており、該当する遺伝子が不明であるもののその脂質関連形質に与える影響は大きなものがあると考えられる。

今後の予定として、これらの多型が実際に遺伝子発現にどのような影響を与えているかを、脂肪前駆細胞 PSPA を用いたレポーターアッセイ等で明らかにするとともに、遺伝型と脂肪関連形質との相関解析による関連性の実証を行うことが重要であると考えられる。また、西洋品種内でも SNP 等の多型は数多く検出されており、その中には転写因子結合に影響を与えることが想定されるものが非常に多く存在する。これらについては西洋品種の改良につながることを期待され、育種のためのマーカーとしての期待は大きいものがあると考えられる。

5. 主な発表論等

(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 0 件)

[学会発表] (計 1 件)

① Uenishi H, Toshimi M, Eguchi-Ogawa T, Nakajima I. Polymorphisms in promoter regions of lipid/adipocyte-related genes in pigs and their distribution among different breeds. International Plant and Animal Genome Conference. 2016. 1. 9 San Diego (USA)

[その他]

ホームページ等

① ブタ発現遺伝子データベース
Pig Expression Data Explorer (PEDE)
<http://pede.dna.affrc.go.jp/>

6. 研究組織

(1) 研究代表者

上西 博英 (UENISHI, Hirohide)

国立研究開発法人農業・食品産業技術総合研究機構・生物機能利用研究部門 動物機能利用研究領域・ユニット長

研究者番号：80391556

(2) 研究分担者

中島 郁世 (NAKAJIMA, Ikuyo)

国立研究開発法人農業・食品産業技術総合研究機構・畜産研究部門 畜産物研究領域・上級研究員

研究者番号：60355063

小川 智子 (OGAWA, Tomoko)

国立研究開発法人農業・食品産業技術総合研究機構・本部 評価室・主任研究員

研究者番号：90466011