

筑波大学

博士（医学）学位論文

機能性 SNP rs7074440 による C-FOS を介した
TCF7L2 遺伝子発現調節機構

2017

筑波大学大学院博士課程人間総合科学研究科

朴 賢英

要旨

目 的:

生活習慣病に関連する多くの Single nucleotide polymorphism (SNP) が Genome-wide association study (GWAS) によって同定されており、その中でも Transcription factor 7-like 2 (TCF7L2) 遺伝子上に存在するリード SNP rs7903146 は 2 型糖尿病発症と最も相関する SNP として広く認知されている。しかし、その SNP がどのようなメカニズムにより遺伝子発現に影響を与え、2 型糖尿病発症の原因となりうるかに関しては未解明のままである。そこで本研究ではリード SNP rs7903146 を含む周辺の SNPs の中に、TCF7L2 遺伝子の発現を直接制御する機能性 SNP が存在すると着想し解析を行った。

対象と方法:

初めに遺伝子発現調節を担うエンハンサーとなりうる機能性 SNP の候補を選定するため、Web 上のデータベース HaploReg V4 を用いてリード SNP rs7903146 と強い連鎖不平衡にある SNP を、SNP rs7903146 を含めて 21 個選出した。選出された各 SNPs の各アレルについて、SNP 周辺 25bp を 3 回タンデムに繰り返した配列を、SV40 プロモーターに連結されたレポーター遺伝子の 3'側に挿入した。レポーターとして各 SNP のリスクアレルはホタルルシフェラーゼ、ノーマルアレルはレニラルシフェラーゼを用いた。TCF7L2 遺伝子は肝臓での糖代謝に重要な役割を担っているとの報告があるため、培養ヒト肝細胞株 HepG2 ならびに Huh7 細胞において各 SNP のノーマルアレルとリスクアレルにおけるエンハンサー活性を、ルシフェラーゼレポーターアッセイを用いて評価した。

結 果:

ルシフェラーゼレポーターアッセイの結果、21 個の候補 SNPs のうち SNP rs7074440 のノーマル (G) アレルがリスク (A) アレルに比べ再現性のある高いレポーター活性を示した。次に SNP rs7074440 のエンハンサー活性を厳密に評価するため、ノーマル (G) アレルとリスク (A) アレル周辺 25bp を 3 回タンデムに繰り返した配列をホタルルシフェラーゼレポーター遺伝子の 3'側にそれぞれ組み込み、同一のレポーターで比較検討を行った。その結果、SNP rs7074440 のノーマル (G) アレルはリスク (A) アレルに比べて、約 15 倍レポーター活性が高いことが示された。引き続き、SNP rs7074440 に結合する転写因子を同定するため、Transcription factor expression library

(TFEL)を用いたゲノムワイドな探索を行った結果、C-FOS が結合因子として見出された。C-FOS の SNP rs7074440 への結合は Electrophoretic mobility shift assay (EMSA)法や、Chromatin immunoprecipitation (ChIP)法によっても確認された。さらに、HepG2 ならびに Huh7 細胞にて C-FOS をノックダウンしたところ、TCF7L2 遺伝子の有意な発現低下が認められた。

考 察:

SNP rs7074440 の Minor allele frequency (MAF) は、ヨーロッパ、北米、アフリカおよびアジア地域の集団においては、それぞれ 0.32、0.22、0.16 および 0.01 であるのに対し、リード SNP rs7903146 の MAF は、0.31、0.25、0.25 および 0.03 である。これら 2 つの SNP の連鎖不平衡 ((Linkage disequilibrium (LD)) 係数は $r^2=0.85$ であり、SNP rs7074440 がリード SNP rs7903146 の機能性 SNP である可能性がある。TCF7L2 に関しては膵 β 細胞においても多くの解析がなされており、マウス培養膵 β 細胞株ではリード SNP rs7903146 自体が TCF7L2 の遺伝子発現に影響を与えるエンハンサーであることが報告されている。しかし、本研究で行った培養ヒト肝細胞における解析では SNP rs7903146 がエンハンサー活性を有することは追認されなかった。これらの結果より SNP のエンハンサー活性は細胞や組織ごとに異なることが示唆される。SNP rs7074440 に結合する C-FOS は JUN とヘテロ二量体を形成して複合体を形成し、特異的な DNA 配列に結合して周辺遺伝子の発現に影響を及ぼす転写因子である。C-FOS を含む Fos ファミリーは一般的に細胞の分化と増殖に関与し、インスリンを含む様々な成長因子によりその転写活性が制御されている。C-FOS および JUN は肝細胞において糖新生遺伝子発現を負に調節すること、さらには、TCF7L2 は肝臓での糖新生を抑制するという報告があるため、本研究で見出された C-FOS による TCF7L2 遺伝子の発現調節メカニズムは、肝臓における糖新生遺伝子発現に深く関係し、2 型糖尿病の発症機序にも関与していることが示唆される。

結 論:

本研究では、2 型糖尿病発症と最も相関のあるリード SNP rs7903146 と連鎖不平衡にある SNP rs7074440 が、肝細胞において転写因子 C-FOS との結合を介して TCF7L2 遺伝子発現を調節している機能性 SNP であることを見出した。

目次

要旨	2
略語一覽	5
序文	7
実験方法	9
結果	17
考察	21
謝辞	24
参考文献	25
図表	31

略語一覽

ANOVA, Analysis of variance

AP-1, Activator protein-1

CEU, Utah residents with Northern and Western European ancestry from the CEPH collection

CCS, Cosmic calf serum

ChIP, Chromatin immunoprecipitation

ChIP-seq, ChIP-sequencing

EMSA, Electrophoretic mobility shift assay

ENCODE, Encyclopedia of DNA elements

FBS, Fetal bovine serum

eQTL, expression Quantitative trait locus

GLP-1, Glucagon-like peptide-1

GWAS, Genome-wide association study

HMG, High mobility group

LD, Linkage disequilibrium

PCR, Polymerase chain reaction

qPCR, quantitative PCR

SEM, Standard error of mean

SNP, Single nucleotide polymorphism

TCF7L2, Transcription factor 7-like 2

TFEL, Transcription factor expression library

UCSC, University of california santa cruz

UTR, Untranslation region

序文

Single-nucleotide polymorphism (SNP) とはゲノム上で 1%以上の頻度で認められる一塩基違いの多型である。その SNP による塩基の違いが個人差と呼ばれるような身体的所見や、様々な疾患の発症リスクと関連していることが明らかとなっている。

2 型糖尿病は、遺伝的要因と環境的要因の両方によって引き起こされる複雑な疾患である [1]。近年、2 型糖尿病に関連する多くの SNPs が Genome-wide association study (GWAS) によって同定されている[2-5]。その中でも Transcription factor 7-like 2 (TCF7L2) 遺伝子に存在する SNPs は異なる民族間の多数の研究によって 2 型糖尿病発症リスクと最も強く関連するマーカーとして報告されており、その中でも SNP rs7903146 が一番代表的なマーカーとして確立されている[6-15]。過去の研究によれば、このリード SNP rs7903146 のオッズ比は 1.7 以上で、北米やヨーロッパにおいては非常に相関が高いことが明らかになっており、この SNP は白人における糖尿病発症リスクの 20-30%程度、またアジア人でも糖尿病発症リスクの 3-5%程度の関連があるとされている[16]。

TCF7L2 遺伝子は第 10 染色体上に位置し、High mobility group (HMG) ボックスを含む転写因子をコードしている[17]。転写因子 TCF7L2 は Wnt シグナル伝達経路に関与し、エンハンサーやプロモーター領域に存在する TCF モチーフとよばれる特定の塩基配列に結合し周辺遺伝子の転写を調節する。2 型糖尿病との関連に先立ち、TCF7L2 遺伝子は癌の発症に関与していることが知られており、Wnt シグナル伝達経路の異常活性化は、結腸直腸癌および他のタイプの腫瘍の発症につながることを示されている[18,19]。また Wnt リガンドは全身のコレステロール代謝、膵 β 細胞の再生やインスリンの分泌、腸管では Glucagon-like Peptide-1 (GLP-1) 産生、肝臓では糖新生に影響を及ぼすことなどが報告されている[20]。

これまでに発表された TCF7L2 遺伝子と 2 型糖尿病を扱った論文の多くが膵臓、特に β 細胞をメインとする報告であり、それらの多くが TCF7L2 遺伝子に存在するリード SNP rs7903146 は膵 β 細胞においてインスリン分泌障害と関連しているという結論に至っている[3-6,9,21]。しかし、2012 年に Boj らが発表した TCF7L2 遺伝子を膵 β 細胞特異的にノックアウトしたマウスの解析結果からは、目立った糖代謝異常は認められなかった[22]。一方、マウス肝臓での TCF7L2 の発現レベルを変化させると糖代謝に大きな影響を与えることが報告されている[22-24]。このように TCF7L2 遺伝子やその SNP が、どの組織で重要な機能を担っているかについては未だ多くの議論がある。また TCF7L2 遺伝子の選択的スプライシングバリエントの発現レベルに関する報告もあり、リード SNP rs7903146 と

の関連を検討する研究が行われていることや[25,26]、各組織での遺伝子発現と SNP を含む染色体領域との相関をデータベース化した **expression Quantitative trait locus (eQTL)**を用いた研究も近年行われている[27,28]。

ゲノムにおける塩基配列の変異はいくつかの異なるメカニズムによってその個体の表現型に影響を及ぼす。主には(1)遺伝子のアミノ酸をコードするエキソン部位の変異によるタンパク質機能の変化、(2)遺伝子のプロモーターやエンハンサー部位における変異による遺伝子発現の変動、(3)遺伝子の **Untranslation region (UTR)** やイントロン部位の変異による mRNA 安定性や選択的スプライシングバリエーションの変化などが挙げられる。今まで報告されている 2 型糖尿病に関連する **TCF7L2** 遺伝子内の **SNPs** は全てイントロンに存在している[6-15]。すなわち、これらの **SNPs** はアミノ酸をコードする領域には存在せずタンパク質の機能には影響を与えない変異であること、またスプライシングサイトとも離れた位置に存在していること、エンハンサーはイントロンを含む遺伝子領域内のゲノム上にも数多く存在することから、**SNP** が **TCF7L2** 遺伝子の発現に影響を及ぼしている可能性が大きいと推測される。

我々の研究室では、マウス転写因子の大多数を網羅した発現ライブラリー **Transcription Factor Expression Library (TFEL)** と、それを用いたエンハンサーとそこに関わるトランスアクチベーターを見出す有用なスクリーニング方法 (**TFEL スキャン**) を独自に開発した[29]。この **TFEL スキャン** を利用することにより、エンハンサーとトランスアクチベーター間の対応を容易に決定できるようになった。そこで本研究では 2 型糖尿病と最も強く関連するリード **SNP rs7903146** と連鎖不平衡にある **SNP** を手掛かりとして、エンハンサー活性のある機能性 **SNP** を見出し、その **SNP** に対応する転写因子を、**TFEL スキャン** を用いて同定することとした。その結果、リード **SNP rs7903146** と強く連鎖するエンハンサー活性のある機能性 **SNP** を新たに特定し、**TFEL スキャン** を用いて対応するトランスアクチベーターを同定した。

実験方法

【遺伝子の表記方法】

ヒト(Homo sapiens) 遺伝子に関する表記は大文字、マウス(Mus musculus) 遺伝子に関する表記は語頭のみ大文字、残りを小文字とした。

【機能性 SNP の候補の選定】

University of California Santa Cruz (UCSC) ゲノムブラウザ (<https://genome.ucsc.edu/>) は、TCF7L2 遺伝子領域のクロマチン状態、SNP の位置や連鎖不平衡=Linkage disequilibrium (LD) マップを検索するためのゲノムワイドマップのソースとして用いた。

SNP rs7903146 は、GWAS により 2 型糖尿病発症と最も強く関連していることが確立されているリード SNP であるが、GWAS はあくまで“マーカー”となる SNP を特定するための解析である。そこでリード SNP rs7903146 と連鎖している SNPs の中から機能性 SNP を特定するため、HaploRegV4 (<http://compbio.mit.edu/HaploReg>) [30,31]を用いてリード SNP rs7903146 との LD 係数がヨーロッパ、北米、アジアの地域において $r^2 > 0.60$ の強く連鎖する SNPs を選択した。さらにそれらの SNPs の中で、上記の 2 つ以上の地域で $r^2 > 0.60$ であり、かつ全ての地域で Minor allele frequency (MAF) ≥ 0.01 の SNP に絞った。その結果、リード SNP rs7903146 を含めて 21 個の SNP が機能性 SNP の候補として選定された。全ての候補 SNPs は、第 10 番染色体上の 114739724 から 114806999 (UCSC ゲノムブラウザ NCBI36/hg18) におよぶ約 67kb 領域内に存在していた。

【プラスミド作製方法】

【機能性 SNP の候補の選定】の中で選定された各 SNP に関して、リード SNP rs7903146 の 2 型糖尿病発症リスクの高いアレルと強く連鎖するアレルをリスクアレルとし、相対的にリスクの低いアレルをノーマルアレルとした。各 SNP のリスクアレルを含めた周辺 25bp 塩基配列を 1 回または 3 回タンドムに繰り返した DNA 断片を、SV40 プロモーターを持つ pGL3-promoter プラスミド (Promega) のホタルルシフェラーゼレポーター遺伝子 (Firefly Luc) の 3' 末端にある BamHI および Sall サイトにライゲーションした (SV40pro-Firefly Luc-SNPx1, SV40pro-Firefly Luc-SNPx3)。各 SNP のノーマ

ルアレルに関しては、リスクアレルと同様に周辺 25bp 塩基配列を 1 回または 3 回タンデムに繰り返した DNA 断片を、ホタルルシフェラーゼレポーター遺伝子がレニラルシフェラーゼレポーター遺伝子 (Renilla Luc) で置換された改変型 pGL3-promoter プラスミド (SV40pro-Renilla Luc) のレニラルシフェラーゼレポーター遺伝子の 3'末端にある BamHI および Sall サイトにライゲーションした (SV40pro-Renilla Luc-SNPx1、SV40pro-Renilla Luc-SNPx3)。上記で用いた配列は、Table 1 に記載されている。

ヒト TCF7L2 プロモーター領域は、下記のプライマーセット

5'-ACGCGTGGCATATCCATCCTAGTGGGAC-3'

5'-AGATCTGAACGGAGTAGTCTGGGAGC-3'

を用いて Polymerase chain reaction (PCR) により増幅し、PCR 産物を pGEM-T-easy プラスミド (Promega) にクローニングした。クローニングされた DNA 断片をベクターから切り出し、pGL3-basic プラスミド (promega) の MluI および BglII サイトにライゲーションした (TCF7L2pro-Firefly Luc)。そして SNP rs7074440 ノーマル (G) アレルまたはリスク (A) アレルの周辺 25bp 塩基配列を、TCF7L2 pro-Firefly Luc のホタルルシフェラーゼレポーター遺伝子の 3'末端にある BamHI および Sall サイトにライゲーションした (TCF7L2pro-Firefly Luc-rs7074440)。マウス c-Fos 発現プラスミドは、TFEL に含まれているものを使用した[27]。

【トランスフェクションならびにルシフェラーゼレポーターアッセイ】

培養ヒト肝細胞株 HepG2 ならびに Huh7 細胞は、25mM グルコース、100U/ml ペニシリンおよび 100µg/ml ストレプトマイシン、10%FBS を含む DMEM (Wako) 中で培養した。培養ヒト腎細胞株 HEK293 ならびに HEK293A 細胞、培養ヒト子宮頸がん細胞株 Hela 細胞は、25mM グルコース、100U/ml ペニシリンおよび 100µg/ml ストレプトマイシン、10%CCS を含む DMEM (Wako) 中で培養した。所定の発現プラスミド、ホタルルシフェラーゼレポータープラスミドとレニラルシフェラーゼレポータープラスミドを、SuperFect トランスフェクション試薬 (Qiagen) を用いて 48well プレートにまいた細胞にトランスフェクションした[32]。

トランスフェクションは、すべての細胞株で 3 連または 4 連で実施した。トランスフェクションされた DNA の総量は空ベクターで調整した。トランスフェクション後の細胞を 100µl Reporter Lysis Buffer (Promega) で溶解し、遠心分離した上清中のホタルならびにレニラルシフェラーゼ活性をルミノメーターで測定した[32]。

【*in vivo* イメージングを用いたルシフェラーゼ活性の測定】

6 週齢の ICR 系雄マウスは日本 SLC 株式会社から購入した。全てのマウスは、14 時間/10 時間の明暗サイクル、温度制御環境下で維持され、標準的な実験動物用エサと水に自由にアクセスできるようにした。また、全てのマウスの取り扱いは法人規程第 50 号国立大学法人筑波大学動物実験取扱規程に従った。

SV40pro-Firefly Luc-SNPx3 から、SV40 プロモーターとルシフェラーゼレポーター遺伝子に連結された SNP rs7074440 ノーマル (G) アレルまたはリスク (A) アレル SNP 配列の 3 回タンデムに繰り返した領域を含む DNA 断片を切り出し、Gateway エントリーベクター pENTR4 (Invitrogen) に挿入した。アデノウイルスプラスミドは、pENTR4 とデスチネーションベクター pAd-PL-DEST プラスミド (Invitrogen) を用いた相同組換えによって作成した (Ad-SV40pro-Firefly Luc-SNPx3)。PacI 処理したアデノウイルスプラスミドを HEK293A 細胞にトランスフェクションした後、組換えアデノウイルスが細胞内で十分に増殖した状態を確認してから、細胞を回収して凍結融解により破碎し、超遠心機を用いた CsCl 勾配遠心分離によって組換えアデノウイルスを収集した[29]。

in vivo イメージングは、2010 年の Takeuchi らの報告に従って行った[33]。アデノウイルスは ICR 系雄マウスに 1×10^8 P.F.U./body の用量で静脈内注射した。3 日後、マウスの腹腔内に PBS に溶解したルシフェリン (7.5mg / ml) を注入し、吸入麻酔下で In Vivo Imaging System (IVISTM, Xenogen) を用いて肝臓の発光を可視化した。さらに厳密な発光定量のために、50mg 肝臓組織試料を 300µl Reporter Lysis Buffer 中にてポリロンでホモジナイズし、遠心分離した上清中のホタルルシフェラーゼ活性をルミノメーターにて測定した。肝臓へのアデノウイルス感染効率を測定するために、上記で遠心分離したペレットをプロテイナーゼ K で消化し、フェノールクロロホルムとエタノール沈殿で DNA を抽出後、試料間で DNA 量を一定に揃え quantitative-PCR (qPCR) を用いて試料中のアデノウイルス DNA 量を定量した[32]。アデノウイルス DNA の定量に用いたプライマーセットは

5'-GTCCGGTTATGTAAACAATCC-3'

5'-ATGAAGAAGTGTTTCGTCTTCG-3'

であり、アデノウイルスに挿入されたホタルルシフェラーゼ遺伝子を増幅する配列となっている。

【ゲノムワイドな転写因子スクリーニング (TFEL スキャン)】

TFEL とは理研マウス完全長 cDNA プロジェクトの中から転写因子と推定される遺伝子 1588 クロオンを収集し、クロオンごとに pcDNA3.1 発現プラスミド (Invitrogen) に組み込んで構築された発現プラスミドライブラリーである[29]。また TFEL を用いたスクリーニング法を TFEL スキャンと命名している[27]。TFEL スキャンは、TFEL 1588 クロオンから 10 クロオンずつプールした約 160 グループでルシフェラーゼレポーターアッセイを行う 1 次スクリーニングと、その 1 次スクリーニングにてレポーター活性に変化のあったグループのクロオンを個別にルシフェラーゼレポーターアッセイを行う 2 次スクリーニングからなる。1 次ならびに 2 次スクリーニングにおいては、48well プレートにまいた HEK293 細胞に SNP rs7074440 の SV40pro-Renilla Luc-SNPx3、SV40pro-Firefly Luc-SNPx3 と TFEL クロオンを同時にトランスフェクションした。トランスフェクションの 24 時間後に細胞を 100µl Reporter Lysis Buffer で溶解し、遠心分離した上清中のホタルならびにレニラルシフェラーゼ活性を測定した。

【Electrophoretic Mobility Shift Assay (EMSA)】

マウス c-Fos ならびに c-Jun のアミノ酸をコードする全長 cDNA を PCR により増幅した。開始コドン側のプライマーには T7 プロモーター-スパーサー-Kozak 配列-HA タグ配列を付加した。使用したプライマーセットは、

c-Fos:5'-GGATCCTAATACGACTCACTATAGGGAACAGCCACCATGTACCCATACGATGTT
CCAGATTACGCTATGATGTCCTCGGGTTTCA-3'

5'-CCATAGAGCCCACCGCATC-3'、

c-Jun:5'-GGATCCTAATACGACTCACTATAGGGAACAGCCACCATGTACCCATACGATGTT
CCAGATTACGCTATGACTGCAAAGATGGAAACGAC-3'

5'-CCATAGAGCCCACCGCATC-3'

である。TNT Quick Systems (Promega) の T7 Master Mix を 8 μ l、0.2 μ l メチオニン(1mM)と 0.1 μ g PCR 産物を混合し、37 $^{\circ}$ C で 90 分間インキュベートし *in vitro* にて各タンパク質を合成した。EMSA は Shikama らの報告に従って行った[32]。用いた DNA プローブの配列は Table2 に記載した。DNA プローブは、Klenow Fragment (GE Healthcare) で 5'オーバーハングを充填することにより[α - 32 P] dCTP (Perkin Elmer) で標識し、Sephadex G-25 (GE Healthcare) カラムで精製した。 32 P 標識された DNA プローブを、pH7.8、50mM KCl、1mM EDTA、5mM MgCl₂、5mM ジチオスレイトール、30 μ g/ml Polydl-dC、0.1% Triton X-100、10mM HEPES を含む緩衝液中にて、*in vitro* にて合成したタンパク質と共に氷上で 30 分間インキュベートし DNA プローブ-タンパク質複合体を形成させた。DNA プローブ-タンパク質複合体は、1xTBE (pH8.3、89mM ホウ酸、2mM EDTA、89mM Tris) 緩衝液中にて 4.6% ポリアクリルアミドゲル上で電気泳動して分離した。

【Chromatin immunoprecipitation (ChIP) アッセイ】

HepG2 細胞からの核抽出物の調製および ChIP アッセイは、Nishi-Tatsumi らの報告に従って行った[35,36]。ChIP アッセイには、抗 c-Fos 抗体 (sc-52-G) (Santa cruz) とノーマルヤギ IgG 抗体 (sc-2061) (Santa cruz) を用いた。各抗体により収集した DNA 試料を、qPCR を用いて評価した。用いた DNA 試料の 1% を input DNA として使用した。SNP rs7074440 領域を含む TCF7L2 遺伝子領域は下記のプライマーセットを用いて増幅した。

5'-GAATTCCTGGACTCAAGCAATC-3'

5'-ACCTCTCTACGCCTCC AGATC-3'

TCF7L2 遺伝子の 100kb 上流を含む領域をコントロールとした。コントロール領域は下記のプライマーセットを用いて増幅した。

5'-ATCACGAGGTCAAGAGATCG-3'

5'-TCAGCCTCTTGATAGCTGG-3'

【RNA interference】

マウス Fos ファミリーに属する c-Fos、FosB、Fosl1、Fosl2 遺伝子のアミノ酸をコードする全長 cDNA は下記のプライマーセットを用いて増幅した。

c-Fos: 5'-GGATCCATGATGTTCTCGGGTTTCAACG-3'

5'-CTCGAGTGACTGCTCACAGGGCCA-3'

FosB: 5'-AAAGGATCCATGTTTCAAGCTTTTCCCGGA-3'

5'-AAACTCGAGTTACAGAGCAAGAAGGGAGGG-3'

Fosl1: 5'- GGATCCATGTACCGAGACTACGGGGAA-3'

5'-CTCGAGTCACAAAGCC AGGAGTGTAGGA-3'、

Fosl2: 5'-GGATCCATGTACCAGGATTATCCCGGGA-3'

5'- CTCGAGAGGGTTACAGGGCTAGAAGTGTGG-3'。

それぞれの遺伝子に対するプライマーセットを用いてマウス肝臓 cDNA をテンプレートとして PCR に よって増幅し、PCR 産物を pGEM-T-easy プラスミドにクローニングした。それぞれの遺伝子の cDNA 断片をプラスミドから切り出し、N 末端 Flag タグ付き発現プラスミド pcDNA3.1 (Invitrogen) の BamHI と XhoI サイトにライゲーションした。(pcDNA3.1-Flag-c-Fos、pcDNA3.1-Flag-FosB、 pcDNA3.1-Flag-Fosl1、pcDNA3.1-Flag-Fosl2)。

それぞれの遺伝子をノックダウンするためのターゲット配列を持つ shRNA 発現プラスミドは、下記 の DNA 断片を用いて pENTR/U6 エントリーベクター (Invitrogen) にクローニングした。

C-FOSi-1:5'-GAGTCTGAGGAGGCCTTCATGTGAAGCCACAGATGGTGAAGGCCTCCTCA
GACTC-3'

C-FOSi-2:5'-CTTCTATGCAGCAGACTGGGATGTGAAGCCACAGATGGTCCCAGTCTGCTG
CATAGAAG-3'

FOSBi:5'-GTGCAACCCACCCTCATCTCTTGTGAAGCCACAGATGGAGAGATGAGGGTGG
GTTGCAC-3'

FOSL1i:5'-GACTTCCTGCAGGCGGAGACTGTGAAGCCACAGATGGGTCTCCGCCT
GCAGGAAGTC-3'

FOSL2i:5'-GGAGAGATGAGCAGCTGTCTTGTGAAGCCACAGATGGAGACAGCTGC
TCATCTCTCC-3'

ターゲット配列は各遺伝子に関してヒトとマウス遺伝子両方をノックダウン出来るように、アミノ酸コード領域内のヒトとマウス共通塩基配列を用いて設計した。Fos ファミリータンパク質のノックダウン効率を調べるために、HEK293 細胞に各マウス Fos ファミリーの Flag タグ付き発現ベクターと FOS ファミリーノックダウンプラスミドを同時にトランスフェクションした。その後 2016 年の Takeuchi ら[29]の報告に従い、トランスフェクションした細胞から回収したタンパク質を用いて抗 FLAG (F3165) 抗体 (Sigma) を用いてウエスタンブロッティングを行い、各ノックダウンプラスミドが効果的にターゲット遺伝子をノックダウンすることを確認した。

HepG2 細胞における各 Fos ファミリーの役割を調べるため、48well プレートの細胞に各 Fos ファミリー遺伝子に対するノックダウンプラスミド、SV40pro-Firefly Luc-rs7074440x3 ノーマル (G) アレルと SV40pro-Renilla Luc プラスミドを同時にトランスフェクションした。トランスフェクションされた DNA の総量は空ベクターで調整した。トランスフェクション後の細胞を 100µl Reporter Lysis Buffer (Promega) で溶解し、遠心分離した上清中のホタルならびにレニラルシフェラーゼ活性をルミノメーターで測定した[30]。

C-FOSi-1、C-FOSi-2 ノックダウンプラスミドは pAd-PL-DEST との Gateway システムを用いた相同組換えによりノックダウン用のアデノウイルスベクターを作製した (Ad-C-FOSi-1 ならびに Ad-C-FOSi-2)。LacZ-shRNA 発現プラスミド (LacZi) とそのアデノウイルスベクター (Ad-LacZi) は 2010 年の Takeuchi らの報告と同じものを用いた[33]。

【RNA の単離および qPCR】

HEK293、HepG2、Huh7 細胞より Sepasol-RNA I Super G (Nakarai tesque) を用いて RNA を抽出し、逆転写試薬キット ReverTra Ace qPCR RT Kit (TOYOBO) を用いて cDNA に逆転写した。QPCR は、ABI 7500 System (Applied Biosystems) により SYBR Green 試薬 KAPA SYBR Fast

qPCR Kit(日本ジェネティクス)を用いて行った。各遺伝子発現の検討に使用したプライマー配列は下記の通りである。

C-FOS: 5'-CTTCAACGCAGACTACGAGG-3'

5'-ATGAAGTTGGCACTGGAGA-3'

FOSB: 5'-TCACCCTCTGCCGAGTCTC-3'

5'-ACGAAGGAACCGGGCATT-3'

FOSL1: 5'-TTCCGAGACTTCGGGGAA-3'

5'-TTGGCACCAGGTGGAAGTT-3'

FOSL2: 5'-ATTATCCCGGGAAGTTTGACA-3'

5'-GTTGATGGTGGGGATGAATG-3'

TCF7L2: 5'-TCGCCTGGCACCGTAGGACA-3'

5'-GGATGCGGAATGCCCGTCGT-3'

Cyclophilin: 5'-GCATACGGGTCCTGGCATCTTGTCC-3'

5'-ATGGTGATCTTCTTGCTGGTCTTGC-3'

【統計解析】

データは平均±Standard error of mean (SEM)として表した。2 群間の差は両側 unpaired Student-t 検定を用いて評価した。3 つ以上の群の差は、Statcel3 (OSM)を使用して Analysis of variance (ANOVA)によって評価した。Tukey-Kramer 法を用いて $p < 0.05$ で統計的に有意とした。

結果

【機能性 SNP の候補の選定】

リード SNP rs7903146 は、多数の GWAS の結果により 2 型糖尿病リスクと最も強く関連するマーカーであることが確認されており、TCF7L2 遺伝子のイントロン 4 に存在している。リード SNP rs7903146 に関連する機能性 SNP を同定するために、リード SNP rs7903146 に対して強く連鎖している 21 個の SNP を選定した (Figure 1A)。SNP 選定の詳細は、実験方法の【機能性 SNP の候補の選定】の項に記載した。21 個の候補 SNP のうち 9 個の SNP がリード SNP rs7903146 の上流約 10kb 内に、他の 11 個の SNP はその下流約 60kb 内に位置し、全ての SNP は TCF7L2 遺伝子のイントロン 4 に位置していた (Figure 1B)。

【SNP rs7074440 はエンハンサー活性を示した】

肝細胞における TCF7L2 遺伝子発現に影響を及ぼす機能性 SNP を探索するために、培養ヒト肝細胞株 HepG2 および Huh7 細胞においてルシフェラーゼレポーターアッセイを行った。実験方法の【プラスミド作製方法】の項に記載した、各 SNP のノーマルアレル周辺塩基を 3 回タンデムに繰り返した配列をレニラルシフェラーゼレポーター遺伝子の 3'末端に挿入したプラスミド (SV40pro-Renilla Luc-SNPx3) と、リスクアレル周辺塩基を 3 回タンデムに繰り返した配列をホタルルシフェラーゼレポーターの 3'末端に挿入したプラスミド (SV40pro-Firefly Luc-SNPx3) を用いて、各 SNP のエンハンサー活性を評価した (Figure 2A)。各 SNP のノーマルアレルレポーター (Renilla Luc) プラスミドとリスクアレルレポーター (Firefly Luc) プラスミドを同じ細胞にトランスフェクションして、ノーマルアレルとリスクアレルのエンハンサー活性を比較した結果、Figure 2B に示すように HepG2 と Huh7 細胞の両方において、SNP rs7074440 のノーマル (G) アレルがリスク (A) アレルと比較して顕著に高いルシフェラーゼ活性を示した。一方、非肝細胞系の培養細胞株であるヒト腎細胞株 HEK293 細胞およびヒト子宮頸がん細胞株 HeLa 細胞では、いずれの SNP プラスミドについてもエンハンサー活性のアレル特異的差異は観察されなかった (Figure 3)。

Figure 3 の検討ではノーマルアレルとリスクアレルで異なるレポーター遺伝子を用いていたため、SNP rs7074440 のノーマル (G) アレルとリスク (A) アレルにおけるエンハンサー活性を厳密に比較

するために、ノーマル(G)アレルとリスク(A)アレル周辺塩基を 3 回タンデムに繰り返した配列をホタルルシフェラーゼレポータープラスミドにそれぞれ組み込み(SV40pro-Firefly Luc-rs7074440x3)、同一のレポーターで比較検討を行った。HepG2 細胞でのトランスフェクションの結果、Figure 4A に示すようにノーマル(G)アレルの SV40pro-Firefly Luc-rs7074440x3 は、リスク(A)アレルと比較して 15 倍高いルシフェラーゼ活性があることから、SNP rs7074440 のノーマル(G)アレルは強いエンハンサー活性を持つことが示された。SNP rs7074440 のノーマル(G)アレル周辺塩基配列を 1 回のみ挿入したレポータープラスミド(SV40pro-Firefly Luc-rs7074440x1)もまた、HepG2 細胞においてリスク(A)アレルレポータープラスミドと比較して有意により高いルシフェラーゼ活性を示した(Figure 5)。さらに HepG2 細胞において、ヒト TCF7L2 遺伝子プロモーターを用いたレポータープラスミド(TCF7L2 pro-Firefly Luc-rs7074440x1)についても同様の結果が得られた(Figure 4B)。また、Figure 4A と 4B の結果から、SNP rs7074440 のリスク(A)アレルにはエンハンサー活性がないことが示された。

*in vivo*での肝臓における SNP rs7074440 のエンハンサー活性を評価するために、*in vivo* Ad-luc 法[25]での実験を行った。SV40 プロモーターおよびホタルルシフェラーゼレポーター遺伝子、ノーマル(G)アレルおよびリスク(A)アレルを含むアデノウイルス(Ad-SV40pro-Luc-rs7074440x3)をマウスの肝臓に感染させ、*in vivo* イメージングと肝臓ホモジナイズ試料の両方を用いたルシフェラーゼレポーター活性を定量した結果から、マウス肝臓においてもノーマル(G)アレルのエンハンサー活性がリスク(A)アレルと比較して有意に高いことが示された(Figure 4C)。これらの結果から、SNP rs7074440 ノーマル(G)アレルが培養肝細胞だけではなく、*in vivo*での肝臓組織においても強いエンハンサー活性を有することを確認できた。

【ゲノムワイドな解析により SNP rs7074440 に結合する転写因子 c-Fos を同定した】

SNP rs7074440 に結合し TCF7L2 遺伝子の発現を調節する可能性のある転写因子を同定するために、マウスの大部分の転写因子を含む 1588 個の転写因子発現ライブラリーTFEL を用いたスクリーニング法である TFEL スキャンを行った。その結果、転写因子 c-Fos がリスク(A)アレルと比較してノーマル(G)アレルレポータープラスミドと比較して強いレポーター活性の促進を示したため、c-Fos が rs7074440 に関するトランスアクチベーターとして同定された(Figure 6A)。さらに SNP rs7074440 における c-Fos の役割を詳細に調べるために、HEK293 細胞に c-Fos 発現プラスミドを

SV40pro-Luc-rs7074440x3 と同時にトランスフェクションし、ルシフェラーゼアッセイを行った。その結果、TFEL スキャンの結果と一致して、c-Fos はノーマル(G)アレルのルシフェラーゼ活性を有意に上昇させた。一方、リスク(A)アレルのルシフェラーゼ活性に有意な差は見られなかった (Figure 6B)。

【C-FOSはrs7074440に直接結合する】

SNP rs7074440のアレルによる塩基配列の違いがどのようにしてタンパク質-DNA結合能を変化させるかを調べるために、SNP rs7074440領域を含む³²P標識されたDNAプローブと*in vitro*にて合成されたc-Fos、c-Junタンパク質を用いてEMSAを行った。その結果、リスク(A)アレルと比較してノーマル(G)アレルを含むDNAプローブにc-Fosとc-Junヘテロ二量体(activator protein-1 (AP-1))タンパク質複合体が強く結合することが示された (Figure 7A)。AP-1が結合するコンセンサス配列はtgaGtcaであり、SNP rs7074440ノーマル(G)アレル周辺配列tgaGtcaと完全に一致している。一方、リスク(A)アレルはtgaAtcaであり塩基配列がAP-1コンセンサス配列とは1塩基異なっている。さらに、*in vivo*でのSNP rs7074440領域への内因性C-FOSタンパク質の結合を確認するため、ChIPアッセイを行った。その結果、HepG2細胞においてSNP rs7074440領域にC-FOSが結合していることが示された (Figure 7B)。なおHepG2およびHuh7細胞ゲノムがSNP rs7074440ノーマル(G)アレルを有することは確認している。

C-FOSがrs7074440に結合することをさらに多面的に検証するために、UCSCゲノムブラウザ上のEncyclopedia of DNA elements (ENCODE) データベースを用いた探索を行った。その結果、ENCODEプロジェクト(2009年2月UCSCゲノムブラウザ (GRCh37/hg19)アセンブリ)の各種転写因子におけるChIP-seqのデータから、SNP rs7074440とFOS結合部位を示すピークが重複していることが判明した (Figure 8)。

これらの結果は、C-FOSがSNP rs7074440のノーマル(G)アレルに直接結合し、リスク(A)アレルへの変異がその結合を弱めることを示唆している。

【C-FOSのノックダウンはTCF7L2遺伝子発現を減少させる】

肝細胞における TCF7L2 遺伝子発現の調節における C-FOS の役割を明らかにするために、HepG2 および Huh7 細胞において内因性 C-FOS をノックダウンさせる実験を行った。Figure 9A および 9B に示すように、C-FOS をターゲットとする 2 種類の shRNA 発現アデノウイルス(Ad-C-FOSi、Ad-C-FOSi-2)による C-FOS のノックダウンは、HepG2 および Huh7 細胞における TCF7L2 遺伝子発現レベルを有意に減少させた。

【HepG2 細胞における FOS ファミリーの SNP rs7074440 に対する役割】

肝細胞と非肝細胞での SNP rs7074440 エンハンサー活性の違いが、それぞれの細胞における C-FOS 遺伝子発現量で説明できるかを調べるために、HEK293、HepG2、Huh7 細胞で FOS ファミリーの遺伝子発現レベルを比較した。FOSL1 に関しては HEK293 に比べ HepG2、Huh7 で発現が高い結果が得られたが、C-FOS を含めて他の FOS ファミリーに関しては HepG2、Huh7 両細胞株で高発現しているという結果は得られなかった(Figure10)。また、内因性に発現している各 FOS ファミリーにおける SNP rs7074440 に対する役割を調べるため、HepG2 細胞を用いて FOS ファミリーのノックダウンプラスミドを用いたルシフェラーゼレポーター実験を行った。その結果、Figure11 で示したように C-FOS をノックダウンしたときにのみ、SNP rs7074440 ノーマル(G)アレルのレポーター活性が有意に低下していた。

以上の結果を合わせると、C-FOS が肝細胞において rs7074440 ノーマル(G)アレルへの結合を介して TCF7L2 発現を調節していることが示唆される。

考察

本研究では、SNP rs7074440 が、肝細胞における転写因子 C-FOS との相互作用を介して TCF7L2 遺伝子発現を変化させる機能性 SNP であることを示した。

初めに GWAS により 2 型糖尿病発症と強く相関のある SNP として確立されているリード SNP rs7903146 と遺伝的に強く連鎖している機能性 SNP を探索し、TCF7L2 遺伝子発現を制御する新規エンハンサー機能性 SNP の候補として rs7074440 を同定した。SNP rs7074440 は、ヒト第 10 番染色体において TCF7L2 遺伝子のイントロン 4 内に存在するリード SNP rs7903146 から約 27kb 距離の下流の位置する数十個の SNP の 1 つである (Figure 1B)。SNP rs7074440 のリスク(A)アレル頻度 (MAF) は、ヨーロッパ、北米、アフリカおよびアジア地域の集団においては、それぞれ 0.32、0.22、0.16 および 0.01 であり、一方、リード SNP rs7903146 におけるリスク(T)アレル頻度 (MAF) は、0.31、0.25、0.25 および 0.03 である。これらの 2 つの SNP の LD は $r^2=0.85$ (ヨーロッパ) であり、SNP rs7074440 がリード SNP rs7903146 の機能性 SNP である可能性は高い。

TCF7L2 遺伝子領域における SNP とエンハンサー活性を調べた過去の報告では、マウス膵 β 細胞株 MIN6 細胞においてはリード SNP rs7903146 自体が TCF7L2 遺伝子発現に影響を及ぼすことが示しているが[21,37]、本研究での培養ヒト肝細胞株における解析ではリード SNP rs7903146 がエンハンサー活性を有することは見出されなかった。またリード SNP rs7903146 と強く連鎖している SNP rs7074440 とは別の SNP rs4132670 が、Huh7 細胞において 1.3 倍高いレベルのエンハンサー活性を示すことも報告されているが[26]、本研究では SNP rs7074440 のような顕著なエンハンサー活性は認められなかった。これらの結果より SNP のエンハンサー活性は細胞や組織、様々な条件により異なってくることが示唆される。

肝臓におけるグルコース代謝の調節ならびに 2 型糖尿病発症における TCF7L2 遺伝子の役割については様々な議論がなされている。いくつかの報告では、TCF7L2 遺伝子のより低い発現が 2 型糖尿病発症リスクを抑えることが示されている。例えば、Boj らの報告では肝臓特異的 Tcf7l2 ノックアウトマウスは、絶食時に肝臓グルコース産生が減少し、高脂肪食においてグルコース代謝が改善することが示されている[22]。一方、Norton らは肝臓での TCF7L2 の過剰発現は肝臓のグルコース産出を有意に減少させ、TCF7L2 のサイレンシングは肝臓のグルコース産出の顕著な増加を誘発し、糖

新生遺伝子の発現の有意な増加を伴うことを報告している[38]。さらに肝臓特異的ドミナントネガティブ TCF7L2トランスジェニックマウスモデルを用いた複数の研究により、TCF7L2 が肝糖新生の抑制において有益な役割を果たすことが示されている[23,24]。Nortonをはじめとする過去のいくつかの報告は、SNP rs7074440 のリスク (A) アレルが肝細胞における TCF7L2 発現を低下させることを示唆する本研究の見解と一致する[23,24,38]。

次にエンハンサー活性を有する機能性 SNP rs7074440 を同定した後、TFEL スキャンを用いてそのトランスアクチベーターである転写因子 C-FOS を見出した。TFEL スキャンは、エンハンサーとトランスアクチベーター間の対応を容易に決定できる方法である[29]。本研究による TFEL スキャンを用いた機能性 SNP に対するトランスアクチベーター同定の成功により、SNP と転写因子の対応を見出す手法「TFEL SNP スキャン」が新たに確立された。

C-FOS はプロトオンコジーンであり、JUN とヘテロ二量体 (AP-1) を形成し、AP-1 コンセンサス配列を持つ DNA に結合して遺伝子の転写活性化をもたらす[39]。C-FOS は FOSB、FOSL1 および FOSL2 からなる FOS ファミリーに属する転写因子である。一般に、FOS ファミリーは細胞分化および増殖において役割を果たし、インスリンを含む様々な成長因子の制御下に置かれている[40]。本研究において、ヒト肝細胞株 HepG2 と Huh7 細胞では SNP rs7074440 のノーマル (G) アレルレポータープラスミドが顕著に高いルシフェラーゼ活性を示したが、非肝細胞系 HEK293 および HeLa 細胞では高いルシフェラーゼ活性は観察されなかった。HepG2 と Huh7 細胞での C-FOS 遺伝子発現量が高いようであれば、その発現量でそのレポーター活性の違いを説明できると考え、HEK293、HepG2、Huh7 細胞で FOS ファミリーの遺伝子発現レベルを比較した。FOSL1 に関しては HEK293 細胞に比べ HepG2、Huh7 細胞で発現が高い結果が得られたが、C-FOS を含めて他の FOS ファミリーに関しては HepG2、Huh7 両細胞株で高発現しているという結果は得られなかった (Figure10)。また、内因性に発現している各 FOS ファミリーにおける SNP rs7074440 に対する役割を調べるため、HepG2 細胞を用いて FOS ファミリーのノックダウンプラスミドを用いたルシフェラーゼレポーター実験を行った。その結果、Figure11 で示したように C-FOS をノックダウンしたときにのみ、SNP rs7074440 ノーマル (G) アレルのレポーター活性が有意に低下していた。これらの結果は、肝細胞においては FOS ファミリーの中でも C-FOS が SNP rs7074440 と相互作用を介して TCF7L2 遺伝子発現を変化させることを示唆している。

C-FOS および JUN は、肝細胞における糖新生遺伝子発現を負に調節する[41]。TCF7L2 が肝臓の糖新生を抑制できることを考えると、本研究で見出された C-FOS による TCF7L2 遺伝子の発現調節メカニズムは、肝臓における糖新生遺伝子発現に深く関係し、2 型糖尿病の発症機序にも関与していることが示唆される。

謝辞

本研究を遂行するにあたり、御指導、御助言を賜りました筑波大学医学医療系内分泌代謝・糖尿病内科 島野仁 教授に深謝いたします。

筑波大学医学医療系 矢作直也 准教授には終始多大なる御協力を賜りました。心より御礼申し上げます。

筑波大学医学医療系 武内 謙憲 助教には実験操作の基礎から結果の解釈、論文作成に至るまできめ細かい御指摘を頂きました。謹んで御礼申し上げます。

また、日頃の実験において多大なる御協力、御支援を頂いた筑波大学医学医療系内分泌代謝・糖尿病内科研究室の方々に心より感謝申し上げます。

本研究を遂行するにあたり、尊い命を捧げてくれた実験動物の冥福をお祈りいたします。

最後になりましたが、私生活ならびに研究室外活動を支えて頂いた親族、先輩、友人、後輩の方々に心より御礼申し上げます。

参考文献

- [1] S. E. Kahn, M. E. Cooper, and S. Del Prato, Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future, *Lancet* 383 (2014) 1068-1083.
- [2] W. T. C. C. Consortium., Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls, *Nature* 447 (2007) 661-678.
- [3] R. Sladek, G. Rocheleau, J. Rung, C. Dina, L. Shen, D. Serre, P. Boutin, D. Vincent, A. Belisle, S. Hadjadj, B. Balkau, B. Heude, G. Charpentier, T. J. Hudson, A. Montpetit, A. V. Pshezhetsky, M. Prentki, B. I. Posner, D. J. Balding, D. Meyre, C. Polychronakos, and P. Froguel, A genome-wide association study identifies novel risk loci for type 2 diabetes, *Nature* 445 (2007) 881-885.
- [4] L. J. Scott, K. L. Mohlke, L. L. Bonnycastle, C. J. Willer, Y. Li, W. L. Duren, M. R. Erdos, H. M. Stringham, P. S. Chines, A. U. Jackson, L. Prokunina-Olsson, C. J. Ding, A. J. Swift, N. Narisu, T. Hu, R. Pruim, R. Xiao, X. Y. Li, K. N. Conneely, N. L. Riebow, A. G. Sprau, M. Tong, P. P. White, K. N. Hetrick, M. W. Barnhart, C. W. Bark, J. L. Goldstein, L. Watkins, F. Xiang, J. Saramies, T. A. Buchanan, R. M. Watanabe, T. T. Valle, L. Kinnunen, G. R. Abecasis, E. W. Pugh, K. F. Doheny, R. N. Bergman, J. Tuomilehto, F. S. Collins, and M. Boehnke, A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants, *Science* 316 (2007) 1341-1345.
- [5] S. F. Grant, G. Thorleifsson, I. Reynisdottir, R. Benediktsson, A. Manolescu, J. Sainz, A. Helgason, H. Stefansson, V. Emilsson, A. Helgadóttir, U. Styrkarsdóttir, K. P. Magnusson, G. B. Walters, E. Palsdóttir, T. Jonsdóttir, T. Gudmundsdóttir, A. Gylfason, J. Saemundsdóttir, R. L. Wilensky, M. P. Reilly, D. J. Rader, Y. Bagger, C. Christiansen, V. Gudnason, G. Sigurdsson, U. Thorsteinsdóttir, J. R. Gulcher, A. Kong, and K. Stefansson, Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes, *Nat Genet* 38 (2006) 320-323.
- [6] J. C. Florez, K. A. Jablonski, N. Bayley, T. I. Pollin, P. I. de Bakker, A. R. Shuldiner, W. C. Knowler, D. M. Nathan, and D. Altshuler, TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program, *N Engl J Med* 355 (2006) 241-250.
- [7] R. Saxena, L. Gianniny, N. P. Burt, V. Lyssenko, C. Giuducci, M. Sjogren, J. C. Florez, P. Almgren, B. Isomaa, M. Orho-Melander, U. Lindblad, M. J. Daly, T. Tuomi, J. N. Hirschhorn, K. G.

Ardlie, L. C. Groop, and D. Altshuler, Common single nucleotide polymorphisms in TCF7L2 are reproducibly associated with type 2 diabetes and reduce the insulin response to glucose in nondiabetic individuals, *Diabetes* 55 (2006) 2890-2895.

[8] E. Zeggini, M. N. Weedon, C. M. Lindgren, T. M. Frayling, K. S. Elliott, H. Lango, N. J. Timpson, J. R. Perry, N. W. Rayner, R. M. Freathy, J. C. Barrett, B. Shields, A. P. Morris, S. Ellard, C. J. Groves, L. W. Harries, J. L. Marchini, K. R. Owen, B. Knight, L. R. Cardon, M. Walker, G. A. Hitman, A. D. Morris, A. S. Doney, M. I. McCarthy, and A. T. Hattersley, Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes, *Science* 316 (2007) 1336-1341.

[9] V. Lyssenko, R. Lupi, P. Marchetti, S. Del Guerra, M. Orho-Melander, P. Almgren, M. Sjogren, C. Ling, K. F. Eriksson, A. L. Lethagen, R. Mancarella, G. Berglund, T. Tuomi, P. Nilsson, S. Del Prato, and L. Groop, Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes, *J Clin Invest* 117 (2007) 2155-2163.

[10] A. Helgason, S. Palsson, G. Thorleifsson, S. F. Grant, V. Emilsson, S. Gunnarsdottir, A. Adeyemo, Y. Chen, G. Chen, I. Reynisdottir, R. Benediktsson, A. Hinney, T. Hansen, G. Andersen, K. Borch-Johnsen, T. Jorgensen, H. Schafer, M. Faruque, A. Doumatey, J. Zhou, R. L. Wilensky, M. P. Reilly, D. J. Rader, Y. Bagger, C. Christiansen, G. Sigurdsson, J. Hebebrand, O. Pedersen, U. Thorsteinsdottir, J. R. Gulcher, A. Kong, C. Rotimi, and K. Stefansson, Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution, *Nat Genet* 39 (2007) 218-225.

[11] S. Cauchi, Y. El Achhab, H. Choquet, C. Dina, F. Kremler, R. Weitgasser, C. Nejjari, W. Patsch, M. Chikri, D. Meyre, and P. Froguel, TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis, *J Mol Med (Berl)* 85 (2007) 777-782.

[12] S. C. Elbein, W. S. Chu, S. K. Das, A. Yao-Borengasser, S. J. Hasstedt, H. Wang, N. Rasouli, and P. A. Kern, Transcription factor 7-like 2 polymorphisms and type 2 diabetes, glucose homeostasis traits and gene expression in US participants of European and African descent, *Diabetologia* 50 (2007) 1621-1630.

- [13] S. Mayans, K. Lackovic, P. Lindgren, K. Ruikka, A. Agren, M. Eliasson, and D. Holmberg, TCF7L2 polymorphisms are associated with type 2 diabetes in northern Sweden, *Eur J Hum Genet* 15 (2007) 342-346.
- [14] M. C. Ng, C. H. Tam, V. K. Lam, W. Y. So, R. C. Ma, and J. C. Chan, Replication and identification of novel variants at TCF7L2 associated with type 2 diabetes in Hong Kong Chinese, *J Clin Endocrinol Metab* 92 (2007) 3733-3737.
- [15] K. Miyake, Y. Horikawa, K. Hara, K. Yasuda, H. Osawa, H. Furuta, Y. Hirota, K. Yamagata, Y. Hinokio, Y. Oka, N. Iwasaki, Y. Iwamoto, Y. Yamada, Y. Seino, H. Maegawa, A. Kashiwagi, K. Yamamoto, K. Tokunaga, J. Takeda, H. Makino, K. Nanjo, T. Kadowaki, and M. Kasuga, Association of TCF7L2 polymorphisms with susceptibility to type 2 diabetes in 4,087 Japanese subjects, *J Hum Genet* 53 (2008) 174-180.
- [16] A. Ramachandran, R.C. Ma, C. Snehalatha. Diabetes in Asia, *Lancet* 375 (2010) 408-418.
- [17] T. Jin. Current Understanding on Role of the Wnt Signaling Pathway Effector TCF7L2 in Glucose Homeostasis, *Endocr Rev* 37 (2016) 254-277.
- [18] M.L. Slattery, A.R. Folsom, R. Wolff, J. Herrick, B.J. Caan, J.D. Potter. Transcription factor 7-like 2 polymorphism and colon cancer, *Cancer Epidemiol Biomarkers Prev* 17 (2008) 978-982.
- [19] A. Saadeddin, R. Babaei-Jadidi, B. Spencer-Dene, A.S. Nateri, The links between transcription, beta-catenin/JNK signaling, and carcinogenesis, *Mol Cancer Res* 8(2009) 1189-1196.
- [20] S.A. Schäfer, F. Machicao, A. Fritsche, H.U. Häring, K. Kantartzis. New type 2 diabetes risk genes provide new insights in insulin secretion mechanisms, *Diabetes Res Clin Pract.* 93 (2011) S9-S24
- [21] K. J. Gaulton, T. Nammo, L. Pasquali, J. M. Simon, P. G. Giresi, M. P. Fogarty, T. M. Panhuis, P. Mieczkowski, A. Secchi, D. Bosco, T. Berney, E. Montanya, K. L. Mohlke, J. D. Lieb, and J. Ferrer, A map of open chromatin in human pancreatic islets, *Nat Genet* 42 (2010) 255-259.
- [22] S. F. Boj, J. H. van Es, M. Huch, V. S. Li, A. Jose, P. Hatzis, M. Mokry, A. Haegebarth, M. van den Born, P. Chambon, P. Voshol, Y. Dor, E. Cuppen, C. Fillat, and H. Clevers, Diabetes risk gene and

Wnt effector Tcf7l2/TCF4 controls hepatic response to perinatal and adult metabolic demand, *Cell* 151 (2012) 1595-1607.

[23] K. J. Oh, J. Park, S. S. Kim, H. Oh, C. S. Choi, and S. H. Koo, TCF7L2 modulates glucose homeostasis by regulating CREB- and FoxO1-dependent transcriptional pathway in the liver, *PLoS Genet* 8 (2012) e1002986.

[24] W. Ip, W. Shao, Z. Song, Z. Chen, M. B. Wheeler, and T. Jin, Liver-specific expression of dominant-negative transcription factor 7-like 2 causes progressive impairment in glucose homeostasis, *Diabetes* 64 (2015) 1923-1932.

[25] B. Neve, O. Le Bacquer, S. Caron, M. Huyvaert, A. Leloire, O. Poulain-Godefroy, C. Lecoeur, F. Pattou, B. Staels, P. Froguel. Alternative human liver transcripts of TCF7L2 bind to the gluconeogenesis regulator HNF4 α at the protein level. *Diabetologia* 57 (2014) 785-96.

[26] D. X. Pang, A. J. Smith, and S. E. Humphries, Functional analysis of TCF7L2 genetic variants associated with type 2 diabetes, *Nutr Metab Cardiovasc Dis* 23 (2013) 550-556.

[27] SC. Elbein, ER. Gamazon, SK. Das, N. Rasouli, PA. Kern, NJ. Cox. Genetic risk factors for type 2 diabetes: a trans-regulatory genetic architecture. *Am J Hum Genet* 91(2012) 466-477.

[28] Y. Zhou. Mechanisms by which variants in the TCF7L2 gene increase the risk of developing Type 2 diabetes, Lund University (2013)

[29] Y. Takeuchi, N. Yahagi, Y. Aita, Y. Murayama, Y. Sawada, X. Piao, N. Toya, Y. Oya, A. Shikama, A. Takarada, Y. Masuda, M. Nishi, M. Kubota, Y. Izumida, T. Yamamoto, M. Sekiya, T. Matsuzaka, Y. Nakagawa, O. Urayama, Y. Kawakami, Y. Iizuka, T. Gotoda, K. Itaka, K. Kataoka, R. Nagai, T. Kadowaki, N. Yamada, Y. Lu, M. Jain, and H. Shimano, KLF15 enables rapid switching between lipogenesis and gluconeogenesis during fasting, *Cell Reports* 16 (2016) 2373-2386.

[30] L. D. Ward, and M. Kellis, HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease, *Nucleic Acids Res* 44 (2016) D877-881.

- [31] R. Saxena, B. F. Voight, V. Lyssenko, N. P. Burtt, P. I. de Bakker, H. Chen, J. J. Roix, S. Kathiresan, J. N. Hirschhorn, M. J. Daly, T. E. Hughes, L. Groop, D. Altshuler, P. Almgren, J. C. Florez, J. Meyer, K. Ardlie, K. Bengtsson Bostrom, B. Isomaa, G. Lettre, U. Lindblad, H. N. Lyon, O. Melander, C. Newton-Cheh, P. Nilsson, M. Orho-Melander, L. Rastam, E. K. Speliotes, M. R. Taskinen, T. Tuomi, C. Guiducci, A. Berglund, J. Carlson, L. Gianniny, R. Hackett, L. Hall, J. Holmkvist, E. Laurila, M. Sjogren, M. Sterner, A. Surti, M. Svensson, R. Tewhey, B. Blumenstiel, M. Parkin, M. Defelice, R. Barry, W. Brodeur, J. Camarata, N. Chia, M. Fava, J. Gibbons, B. Handsaker, C. Healy, K. Nguyen, C. Gates, C. Sougnez, D. Gage, M. Nizzari, S. B. Gabriel, G. W. Chirn, Q. Ma, H. Parikh, D. Richardson, D. Ricke, and S. Purcell, Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels, *Science* 316 (2007) 1331-1336.
- [32] A. Shikama, H. Shinozaki, Y. Takeuchi, T. Matsuzaka, Y. Aita, T. Murayama, Y. Sawada, X. Piao, N. Toya, Y. Oya, A. Takarada, Y. Masuda, M. Nishi, M. Kubota, Y. Izumida, Y. Nakagawa, H. Iwasaki, K. Kobayashi, S. Yatoh, H. Suzuki, H. Yagyu, Y. Kawakami, N. Yamada, H. Shimano, and N. Yahagi, Identification of human ELOVL5 enhancer regions controlled by SREBP, *Biochem Biophys Res Commun* 465 (2015) 857-863.
- [33] Y. Takeuchi, N. Yahagi, Y. Izumida, M. Nishi, M. Kubota, Y. Teraoka, T. Yamamoto, T. Matsuzaka, Y. Nakagawa, M. Sekiya, Y. Iizuka, K. Ohashi, J. Osuga, T. Gotoda, S. Ishibashi, K. Itaka, K. Kataoka, R. Nagai, N. Yamada, T. Kadowaki, and H. Shimano, Polyunsaturated fatty acids selectively suppress sterol regulatory element-binding protein-1 through proteolytic processing and autoloop regulatory circuit, *J Biol Chem* 285 (2010) 11681-11691.
- [34] Y. Takeuchi, N. Yahagi, Y. Nakagawa, T. Matsuzaka, R. Shimizu, M. Sekiya, Y. Iizuka, K. Ohashi, T. Gotoda, M. Yamamoto, R. Nagai, T. Kadowaki, N. Yamada, J. Osuga, and H. Shimano, In vivo promoter analysis on refeeding response of hepatic sterol regulatory element-binding protein-1c expression, *Biochem Biophys Res Commun* 363 (2007) 329-335.
- [35] M. Nishi-Tatsumi, N. Yahagi, Y. Takeuchi, N. Toya, A. Takarada, Y. Murayama, Y. Aita, Y. Sawada, X. Piao, Y. Oya, A. Shikama, Y. Masuda, M. Kubota, Y. Izumida, T. Matsuzaka, Y. Nakagawa, M. Sekiya, Y. Iizuka, Y. Kawakami, T. Kadowaki, N. Yamada, and H. Shimano, A key

role of nuclear factor Y in the refeeding response of fatty acid synthase in adipocytes, *FEBS Lett* 591 (2017) 965-978.

[36] M. Sekiya, N. Yahagi, T. Matsuzaka, Y. Takeuchi, Y. Nakagawa, H. Takahashi, H. Okazaki, Y. Iizuka, K. Ohashi, T. Gotoda, S. Ishibashi, R. Nagai, T. Yamazaki, T. Kadowaki, N. Yamada, J. Osuga, and H. Shimano, SREBP-1-independent regulation of lipogenic gene expression in adipocytes, *J Lipid Res* 48 (2007) 1581-1591.

[37] M. L. Stitzel, P. Sethupathy, D. S. Pearson, P. S. Chines, L. Song, M. R. Erdos, R. Welch, S. C. Parker, A. P. Boyle, L. J. Scott, E. H. Margulies, M. Boehnke, T. S. Furey, G. E. Crawford, and F. S. Collins, Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci, *Cell Metab* 12 (2010) 443-455.

[38] L. Norton, M. Fourcaudot, M. A. Abdul-Ghani, D. Winnier, F. F. Mehta, C. P. Jenkinson, and R. A. Defronzo, Chromatin occupancy of transcription factor 7-like 2 (TCF7L2) and its role in hepatic glucose metabolism, *Diabetologia* 54 (2011) 3132-3142.

[39] R. Chiu, W. J. Boyle, J. Meek, T. Smeal, T. Hunter, and M. Karin, The c-Fos protein interacts with c-Jun/AP-1 to stimulate transcription of AP-1 responsive genes, *Cell* 54 (1988) 541-552.

[40] P. J. Blakeshear, D. M. Haupt, and D. J. Stumpo, Insulin activation of protein kinase C: a reassessment, *J Biol Chem* 266 (1991) 10946-10952.

[41] A. L. Gurney, E. A. Park, M. Giralt, J. Liu, and R. W. Hanson, Opposing actions of Fos and Jun on transcription of the phosphoenolpyruvate carboxykinase (GTP) gene. Dominant negative regulation by Fos, *J Biol Chem* 267 (1992) 18133-18139.

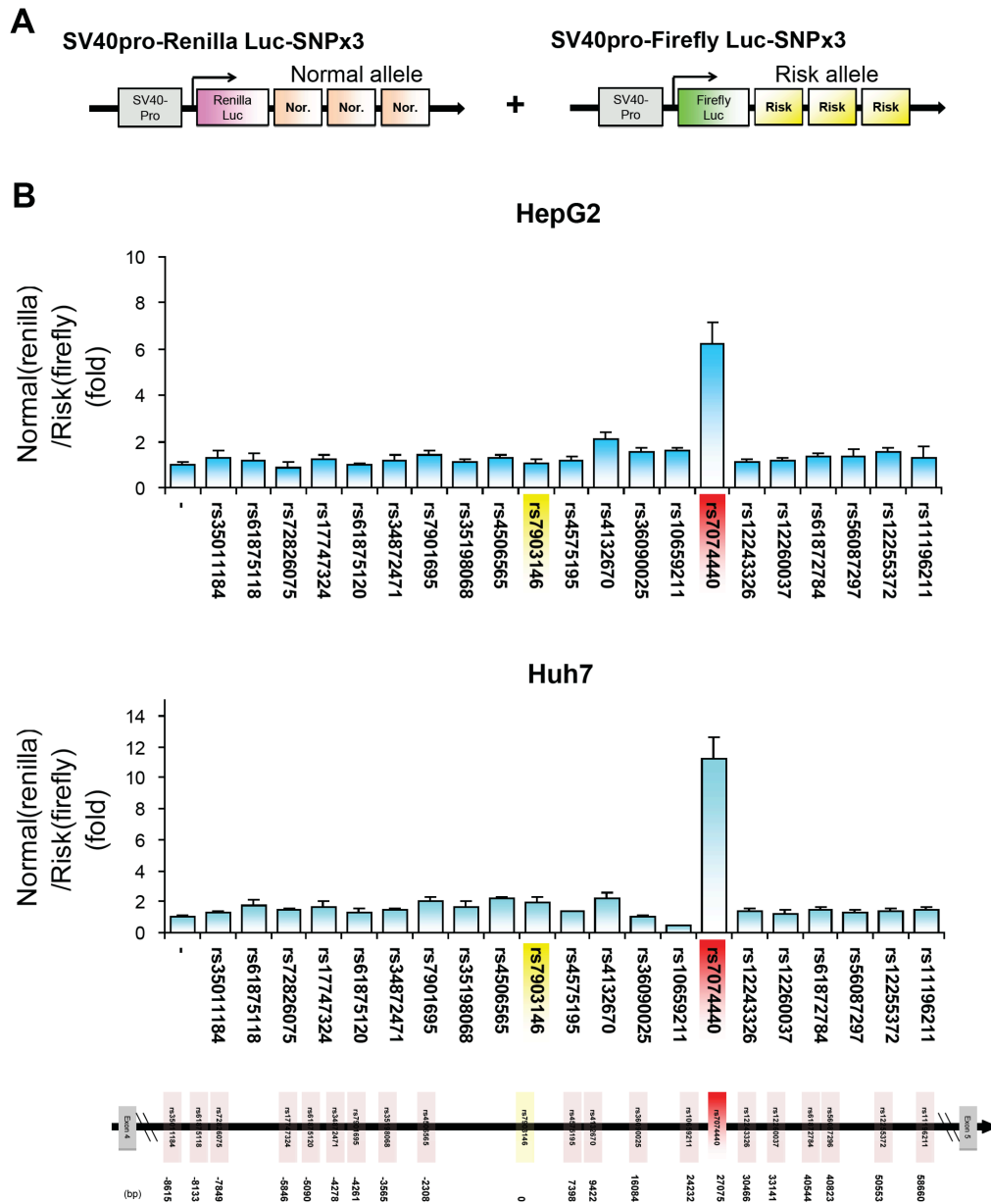


Figure 2. 培養ヒト肝細胞株において SNP rs7074440 ノーマル (G) アレルはリスク (A) アレルに比べ顕著なエンハンサー活性を示した

A) 用いたルシフェラーゼレポータープラスミドのコンストラクト。Figure1で選定した各 SNP 周辺塩基 25bp を 3 回タンデムに繰り返した配列を、SV40 プロモーターを持つルシフェラーゼレポーター遺伝子の 3' 側に挿入してある。各 SNP の「ノーマルアレル」配列をレニナルルシフェラーゼレポーター (SV40pro-Renilla Luc -SNPx3)、「リスクアレル」配列をホタルルシフェラーゼレポーター (SV40pro-Firefly Luc- SNPx3) に挿入した。

B) 各 SNP のノーマルアレル SV40pro-Renilla Luc-SNPx3 とリスクアレル SV40pro-Firefly Luc-SNPx3 をヒト肝細胞 HepG2 および Huh7 細胞にそれぞれトランスフェクションした。縦軸はノーマルアレルとリスクアレルのレポーター活性の比であり、SV40pro のみの群を 1 としている。(n=3)

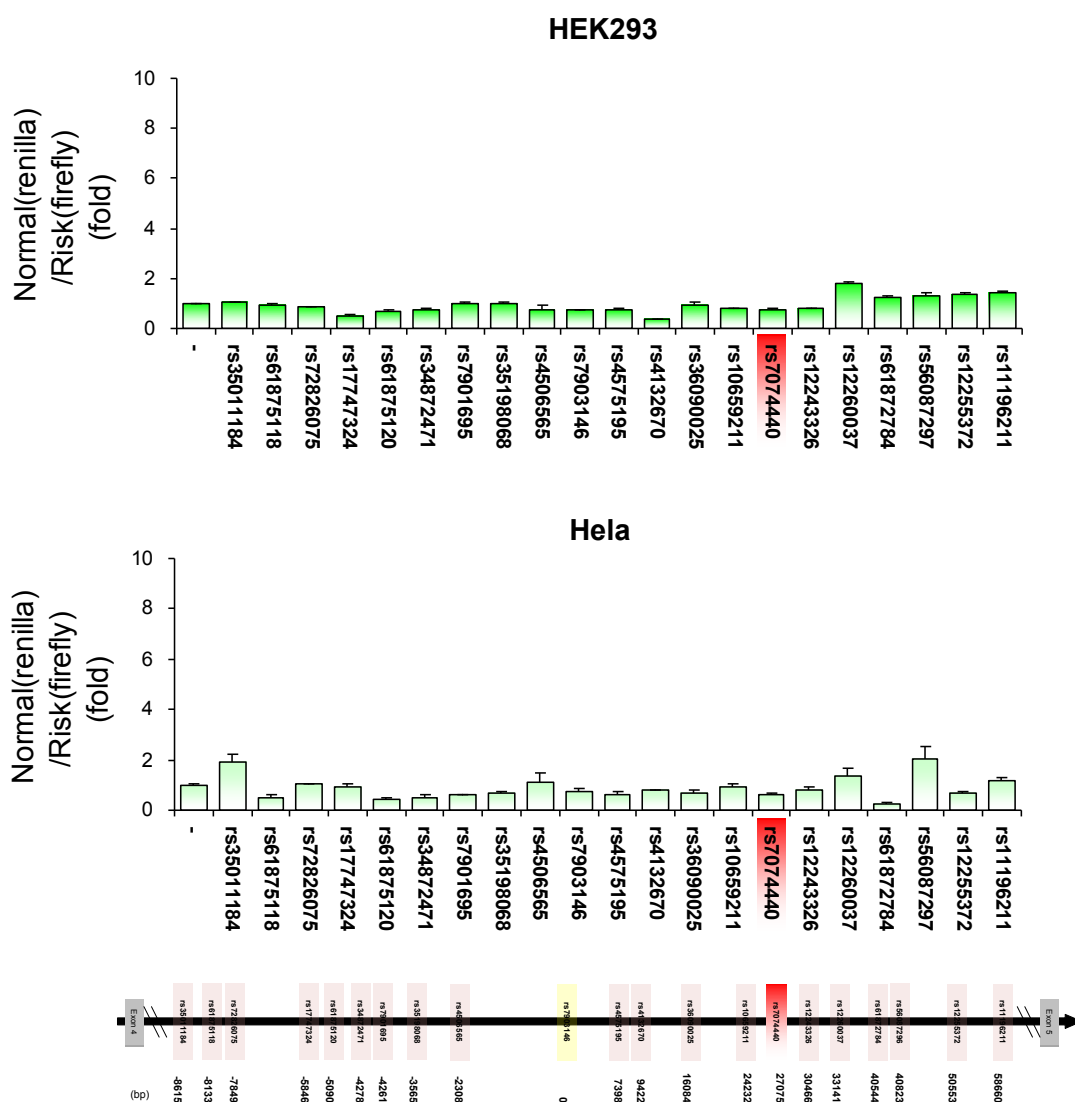


Figure 3. ヒト非肝細胞株においては各 SNP のノーモラルアレルとリスクアレルのエンハンサー活性に顕著な差は見られなかった

各 SNP のノーモラルアレル SV40pro-Renilla Luc-SNPx3 とリスクアレル SV40pro-Firefly Luc-SNPx3 をヒト腎細胞株 HEK293 およびヒト子宮頸がん細胞株 HeLa にそれぞれトランスフェクションした。縦軸はノーモラルアレルとリスクアレルのレポーター活性の比であり、SV40pro のみの群を 1 としている。(n=3)

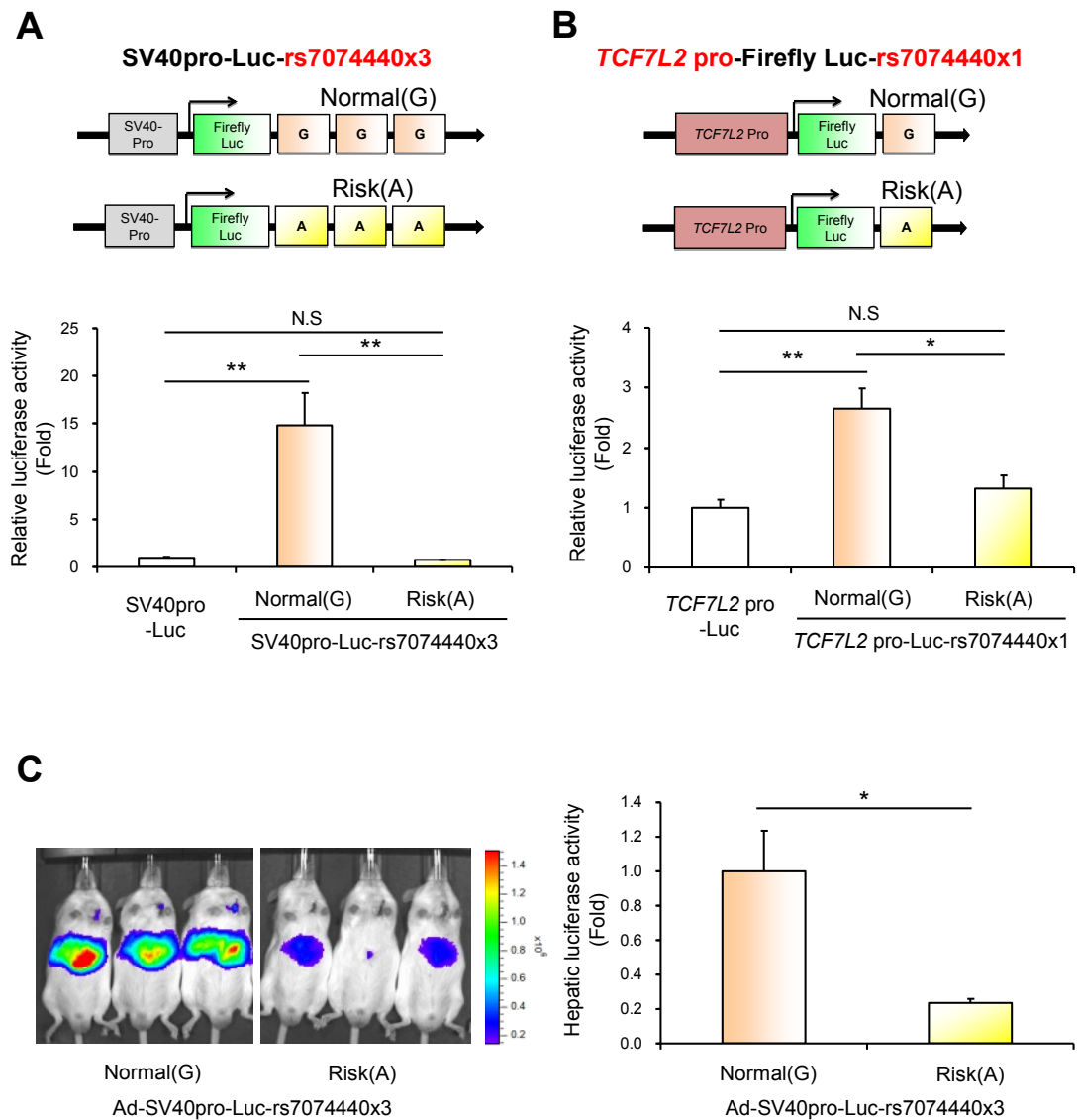


Figure 4. 詳細な SNP rs7074440 のエンハンサー活性の評価

A) SNP rs7074440 ノーマル (G) アレルまたはリスク (A) アレルの周辺塩基 25bp を 3 回 タンデムに繰り返した配列を挿入した SV40 プロモーター-ホタルルシフェラーゼレポーター ベクター (SV40pro-Luc-rs7074440x3) を HepG2 細胞にトランスフェクションした。各群共に SVpro-Renilla Luc を同時にトランスフェクションし、レニラルシフェラーゼ活性で補正している。縦軸は各群のルシフェラーゼ活性を示しており SV40pro のみの群を 1 としている。(n=3)

B) SNP rs7074440 ノーマル (G) アレルまたはリスク (A) アレルの周辺塩基配列 25bp を挿入したヒト TCF7L2 プロモーター-ホタルルシフェラーゼレポーターベクター (TCF7L2 pro-Luc-rs7074440) を HepG2 細胞にトランスフェクションした。各群共に SVpro-Renilla Luc

を同時にトランスフェクションしレニラルシフェラーゼ活性で補正している。縦軸は各群のルシフェラーゼ活性を示して TCF7L2pro のみの群を 1 としている。(n=3)

C) *in vivo* Ad-Luc 法を用いてマウス肝臓における SNP rs7074440 のエンハンサー活性を測定した。SNP rs7074440 ノーマル (G) アレルまたはリスク (A) アレルをホタルルシフェラーゼ遺伝子につないだアデノウイルス (Ad-SV40pro-Luc-rs7074440x3) を ICR 系雄マウスに 1.0×10^8 P.F.U./body にて静脈注射した。3 日後、ルシフェリンを腹腔内注射し、吸入麻酔下で *In Vivo Imaging System* (IVISTM、Xenogen) を用いて肝臓の発光を可視化した (左パネル)。また肝臓組織試料を Reporter Lysis Buffer 中にてホモジナイズし、遠心分離した上清中のルシフェラーゼ活性をルミノメーターにて測定した。ルシフェラーゼ活性は肝臓へ感染したアデノウイルス量で補正している (右パネル)。(n=4-5) * p < 0.05、** P < 0.01、N.S = 有意差無し

SV40pro-Luc-rs7074440x1

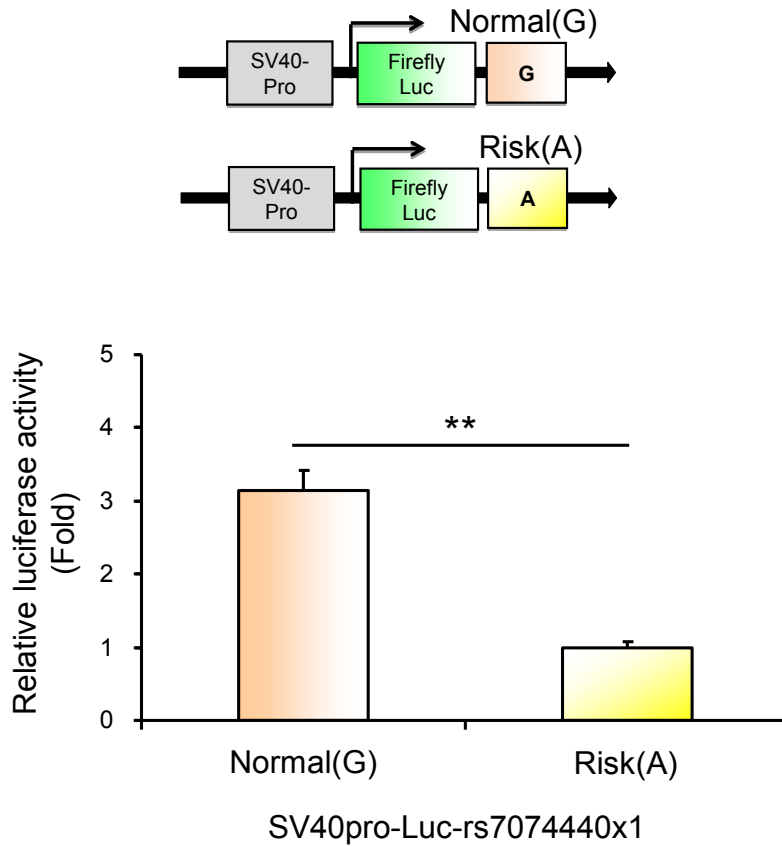


Figure 5. 詳細な SNP rs7074440 のエンハンサー活性の評価

A) SNP rs7074440 ノーマル (G) アレルまたはリスク (A) アレルの周辺塩基配列 25bp を挿入した SV40 プロモーターホタルルシフェラーゼレポーターベクター (SV40pro-Luc-rs7074440x1) を HepG2 細胞にトランスフェクションした。各群共に SVpro-Renilla Luc を同時にトランスフェクションし、レニラルシフェラーゼレポーター活性で補正している。縦軸はリスク (A) アレル群を 1 としている。(n=3) ** P < 0.01。

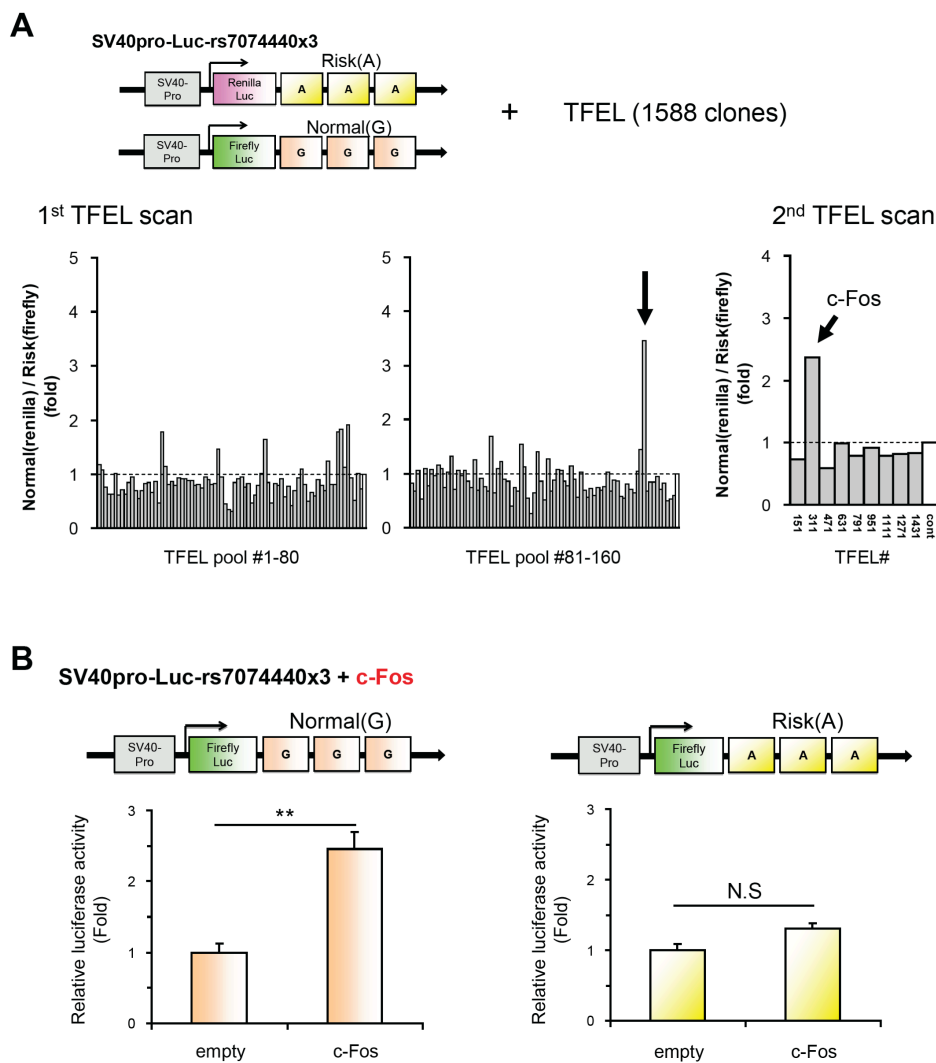


Figure 6. ゲノムワイドな解析により **c-Fos** が **SNP rs7074440** ノーマルアレルのエンハンサー活性を上昇させることを見出した

A) 転写因子ライブラリーTFEL を用いたゲノムワイドスクリーニング (TFEL scan) によって SNP rs7074440 のノーマル (G) アレルとリスク (A) アレルに結合する転写因子を「実験方法の【ゲノムワイドな転写因子スクリーニング (TFEL スキャン)】」に従って探索した。SNP rs7074440 のノーマルアレル SV40pro-Renilla Luc-SNPx3 とリスクアレル SV40pro-Firefly Luc- SNPx3 を用いて、HEK293 細胞にて TFEL スキャン 1 次スクリーニングを行った (左パネル)。1 次でノーマルアレルとリスクアレルのレポーター活性の比が高値であったプールのクローンを、2 次スクリーニングにてクローンごとに検討した (右パネル)。

B) SNP rs7074440 に対する c-Fos の効果を評価した。SNP rs7074440 ノーマル (G) アレル (G) またはリスク (A) アレルの周辺塩基 25bp を 3 回タンデムに繰り返した配列を挿入した SV40 プロモーター-ホタルルシフェラーゼレポーターベクター (SV40pro-Luc-rs7074440 x3) と c-Fos 発現プラスミドを同時にトランスフェクションした。各群共に SVpro-Renilla Luc

を同時にトランスフェクションし、レニラルシフェラーゼ活性で補正している。縦軸は各群の空ベクター群を 1 としている。(n=3) ** P < 0.01、N.S = 有意差無し。

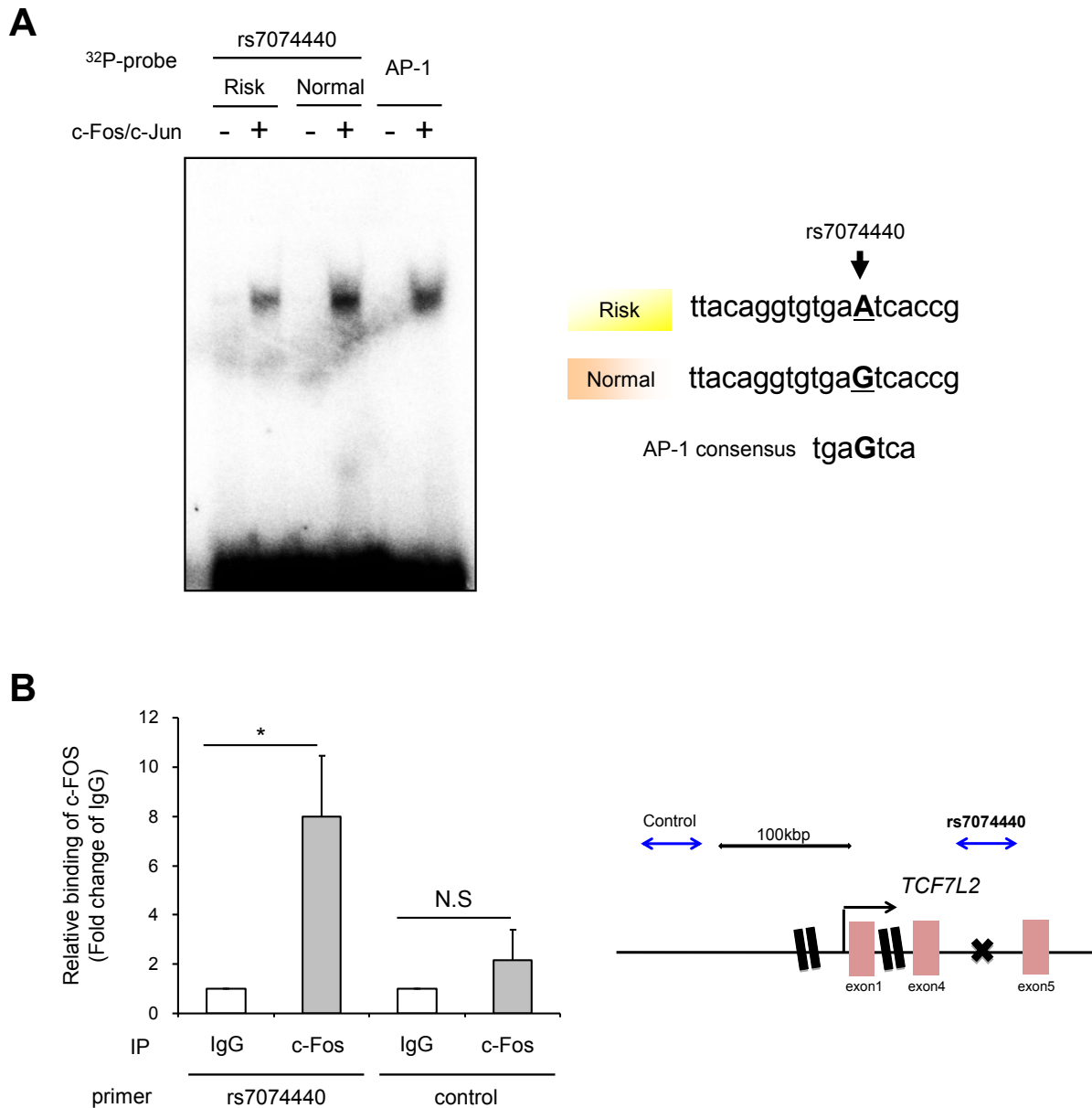


Figure 7. 内因性 C-FOS タンパク質は SNP rs7074440 に結合する。

A) SNP rs7074440 周辺配列および AP-1 コンセンサス配列についての DNA プロブによる EMSA。c-Fos および c-Jun タンパク質を ³²P 標識された DNA プロブとともにインキュベートした。ノーマル (G) アレルは、矢印で示される AP-1 コンセンサス配列と同じ塩基配列を有する。

B) HepG2 細胞における抗 C-FOS 抗体を用いた SNP rs7074440 領域の ChIP アッセイ。抗 C-FOS 抗体により収集した DNA 試料を、SNP rs7074440 を増幅するプライマーセットを用いた qPCR により定量した。また TCF7L2 遺伝子から約 100kb 離れた領域を増幅するプライマー

セットをコントロールとした。縦軸は各群の IgG 抗体による定量結果を 1 としている。(n=3)

* p < 0.05、N.S = 有意差無し。

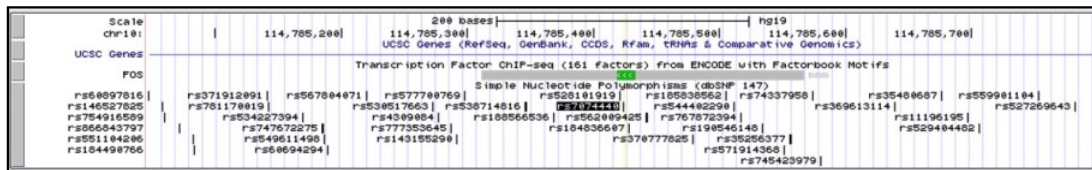
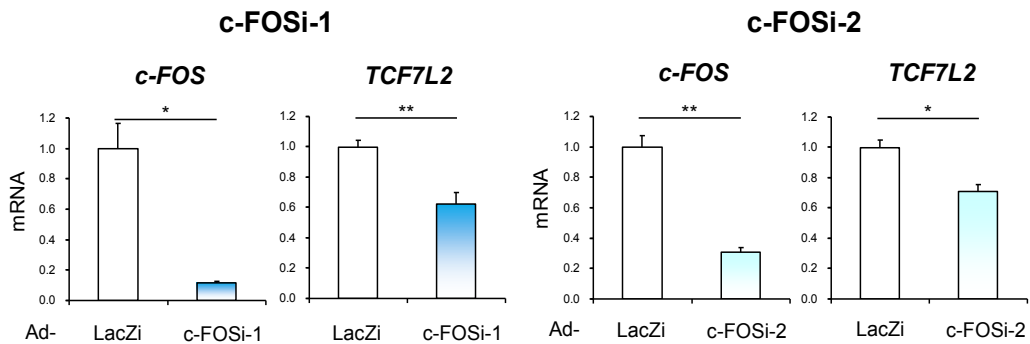


Figure 8. SNP rs7074440 領域における ChIP-seq データ

ゲノムブラウザ ENCODE プロジェクト (GRCh37/hg19) から取得した様々な抗体を用いた ChIP-seq データ。SNP rs7074440 と C-FOS 転写因子結合部位を示すピークが重複している。

A HepG2



B Huh7

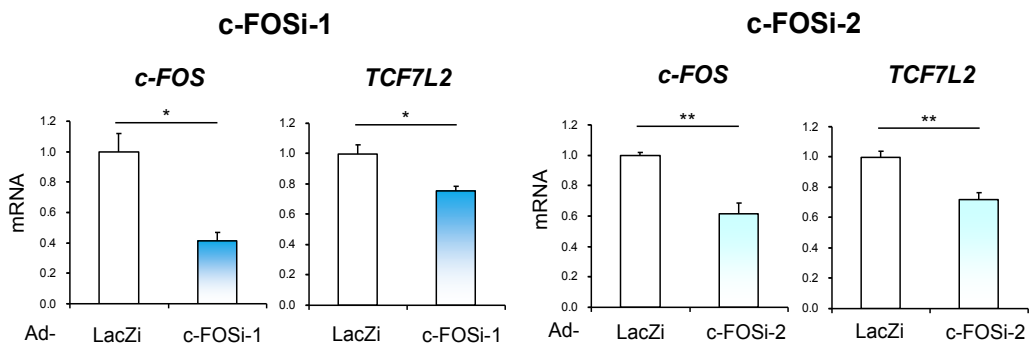


Figure 9. C-FOS のノックダウンは TCF7L2 遺伝子発現を減少させる

C-FOS コード配列に対する shRNA を発現するアデノウイルス (Ad-C-FOSi-1 および Ad-C-FOSi-2) をヒト肝細胞株 HepG2 (A) および Huh7 (B) 細胞に感染させた。C-FOSi-1 と C-FOSi-2 の標的配列は異なる。Ad-LacZi は対照群として用いた。アデノウイルスは細胞に対して 60m.o.i で感染させ、96 時間後に細胞から RNA を採取し、qPCR により各 mRNA を定量した。縦軸は各 mRNA を Cyclophilin mRNA にて補正した値であり、LacZi 群を 1 としている。(n=3) * p < 0.05、** p < 0.01

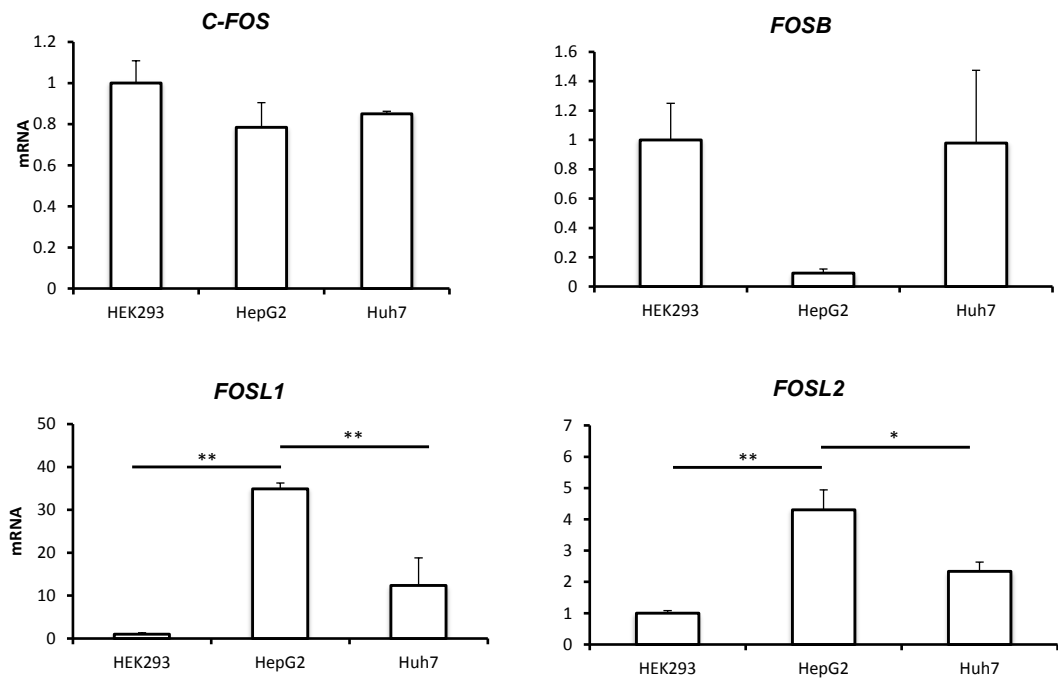


Figure 10. 各細胞における FOS ファミリーの発現量

ヒト腎細胞株 HEK293、ヒト肝細胞株 HepG2 および Huh7 細胞より RNA を採取し、qPCR により FOS ファミリーの各 mRNA を定量した。縦軸は各 mRNA を Cyclophilin mRNA にて補正した値であり、HEK293 群を 1 としている。(n=3) * p<0.05, ** p < 0.01。

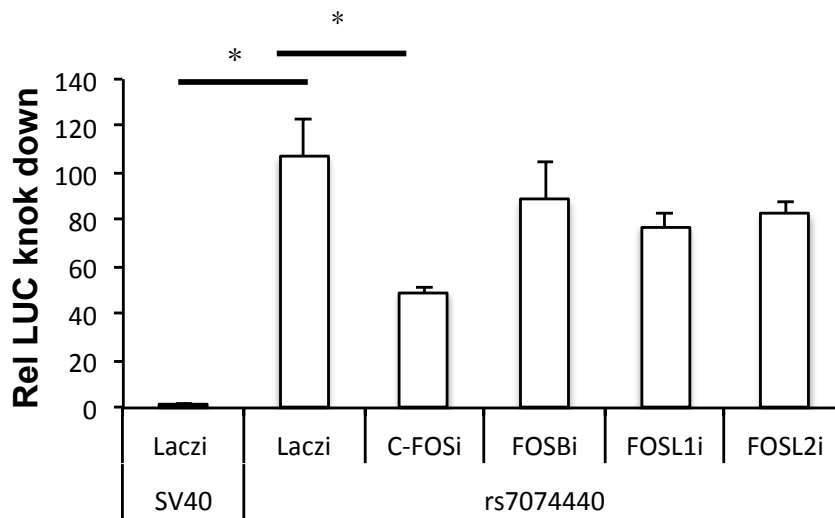


Figure 11. HepG2 細胞における FOS ファミリーをノックダウンさせた際の SNP rs7074440 のエンハンサー活性

HepG2細胞における SNP rs7074440 に対する内因性 FOS ファミリーの効果を評価した。SNP rs7074440 ノーマル (G) アレルの周辺塩基 25bp を 3 回タンデムに繰り返した配列を挿入した SV40 プロモーターホタルルシフェラーゼレポーターベクター (SV40pro-Luc-rs7074440x3) と各 FOS ファミリーノックダウンプラスミドを同時にトランスフェクションした。C-FOSi は C-FOSi-2 を用いた。各群共に SVpro-Renilla Luc を同時にトランスフェクションし、レニラルシフェラーゼ活性で補正している。縦軸は LacZi 群を 1 としている。(n=4) *p <0.05。

rs4575195 Risk allele antisense	TCGACACATTCTATTGTAATGTTCTTTTTCACATTCTATTGTAATGTTCTTTTTCACATTCCCT ATTGTAATGTTCTTTTCG
rs12244851 Normal allele sense	GATCCTGGGTTAGTTCCAGGTCGGCATTCTGGGTTAGTTCACAGGTCGGCATTCTGGGTT AGTTCCAGGTCGGCATTCTG
rs12244851 Normal allele antisense	TCGACGAATGCCGACCTGGGAACTAACCCAGAATGCCGACCTGGGAACTAACCCAGAATG CCGACCTGGGAACTAACCCAG
rs12244851 Risk allele sense	GATCCTGGGTTAGTTCTTAGGTCGGCATTCTGGGTTAGTTCACAGGTCGGCATTCTGGGTT AGTTCTAGGTCGGCATTCTG
rs12244851 Risk allele antisense	TCGACGAATGCCGACCTAGGAACTAACCCAGAATGCCGACCTAGGAACTAACCCAGAATG CCGACCTAGGAACTAACCCAG
rs36090025 Normal allele sense	GATCCGTAATGCCACCTACTCGGGAGGCTGGTAATGCCACCTACTCGGGAGGCTGGTAAT GCCACCTACTCGGGAGGCTGG
rs36090025 Normal allele antisense	TCGACGAGCTCCCGAGTGGCATTACAGCCTCCCGAGTAGGTGGCATTACCAGCC TCCCGAGTAGGTGGCATTACG
rs36090025 Risk allele sense	GATCCGTAATGCCACCTCTCGGGAGGCTGGTAATGCCACCTCTCGGGAGGCTGGTAAT GCCACCTCTCGGGAGGCTGG
rs36090025 Risk allele antisense	TCGACGAGCTCCCGAGGAGGTGGCATTACCAGCCTCCCGAGGAGGTGGCATTACCAGCC TCCCGAGGAGGTGGCATTACG
rs7074440 Normal allele sense	GATCCTTACAGGTGTGAGTACCCGCTCTGTACAGGTGTGAGTACCAGCTCTGTACAG GTGTGAGTACCAGCTCTGG
rs7074440 Normal allele antisense	TCGACGAGGAGCGGTGACTCACACCTGTAAACAGGAGCGGTGACTCACACCTGTAAACAGGA GCGGTGACTCACACCTGTAAAG
rs7074440 Risk allele sense	GATCCTTACAGGTGTGAAATACCAGCTCTGTACAGGTGTGAAATACCAGCTCTGTACAG GTGTGAAATACCAGCTCTGG
rs7074440 Risk allele antisense	TCGACGAGGAGCGGTGATTACACCTGTAAACAGGAGCGGTGATTACACCTGTAAACAGGA GCGGTGATTACACCTGTAAAG
rs10659211 Normal allele sense	GATCCAGCTGTGCTGAGAGGTTGAGTTCAGCTGTGCTGAGAGGTTGAGTTCAGCTGTG CTGAGAGGTTGAGTTCG
rs10659211 Normal allele antisense	TCGACGAACTACCTCTCAGCACAGCTGAACTACCTCTCAGCACAGCTGAACTCAC CTCTCTCAGCACAGCTG
rs10659211 Risk allele sense	GATCCAGCTGTGCTGAGCTAGAGGTGAGTTCAGCTGTGCTGAGCTAGAGGTGAGTTCAGC TGTGCTGAGCTAGAGGTGAGTTCG
rs10659211 Risk allele antisense	TCGACGAACTACCTCTAGCTCAGCACAGCTGAACTACCTCTAGCTCAGCACAGCTGAACT TCACCTCTAGCTCAGCACAGCTG
rs35011184 Normal allele sense	GATCCCTTTCATTGAGCGCTTCTGCGAATGCTTTCATTGAGCGCTTCTGCGAATGCTTTC TTGAGCGCTTCTGCGAATG
rs35011184 Normal allele antisense	TCGACCATTCGCAGAGCGCTCAATGAAAGCAATTCGCAGAGCGCTCAATGAAAGCATT GCAGAGCGCTCAATGAAAGG
rs35011184 Risk allele sense	GATCCCTTTCATTGAGCACTTCTGCGAATGCTTTCATTGAGCACTTCTGCGAATGCTTTC ATTGAGCACTTCTGCGAATG
rs35011184 Risk allele antisense	TCGACCATTCGCAGAGTCTCAATGAAAGCAATTCGCAGAGTCTCAATGAAAGCATT CAGAAAGTCTCAATGAAAGG
rs72826075 Normal allele sense	GATCCAAAAATAAAAAAATTAGCTAGCCAAAAATAAAAAAATTAGCTAGCCAAAAAAT AAAAAAAATTAGCTAGCCAG
rs72826075 Normal allele antisense	TCGACTGGCTAGCTAATTTTTTATTTTTTGGCTAGCTAATTTTTTATTTTTTGGCTAGC TAATTTTTTATTTTTT
rs72826075 Risk allele sense	GATCCAAAAATAAAAAAATTAGCTAGCCAAAAATAAAAAAATTAGCTAGCCAAAAAAT AAAAAAGATTAGCTAGCCAG
rs72826075 Risk allele antisense	TCGACTGGCTAGCTAATTTTTTATTTTTTGGCTAGCTAATTTTTTATTTTTTGGCTAG CTAATTTTTTATTTTTT
rs61875118 Normal allele sense	GATCCGGGAGGCTGAGGGGGTGGATCGCTGGGAGGCTGAGGCGGGTGGATCGCTGGGA GGCTGAGGGGGTGGATCGCTG
rs61875118 Normal allele antisense	TCGACAGCGATCCACCCGCTCAGCCTCCAGCGATCCACCCGCTCAGCCTCCAGCGGA TCCACCCGCTCAGCCTCCCG
rs61875118 Risk allele sense	GATCCGGGAGGCTGAGGTGGTGGATCGCTGGGAGGCTGAGGTGGTGGATCGCTGGGA GGCTGAGGTGGTGGATCGCTG
rs61875118 Risk allele antisense	AGCGATCCACCCACCTCAGCCTCCAGCGATCCACCCACCTCAGCCTCCAGCGATCCACC CACCTCAGCCTCCCG
rs4267006 Normal allele sense	GATCCCTATAAAAAATTATTTTTTCTATTAATAAAAAATTATTTTTTCTATTAATA AAAAATTATTTTTT
rs4267006 Normal allele antisense	TCGACAAAAATAAAAAAATTATTTTTTAAATAAAAAATAAAAAAATTATTTTTTAAATA AAATAATTTTTTAAATAGG
rs4267006 Risk allele sense	GATCCCTATAAAAAATTATTTTTTCTATTAATAAAAAATTATTTTTTCTATTAATA AAAAATTATTTTTT
rs4267006 Risk allele antisense	TCGACAAAAATAAAAAAATTATTTTTTAAATAAAAAATAAAAAAATTATTTTTTAAATA AAATACTTTTTTAAATAGG
rs12260037 Normal allele sense	GATCCCTCAAAGTGATCTGCCACCTCTGCCTCAAAGTGATCTGCCACCTCTGCCTCAA GTGATCTGCCACCTCTGCG
rs12260037 Normal allele antisense	TCGACGAGAGGTGGGCGATCACTTTGAG GCAGAGGTGGGAGATCACTTTGAGGAGAGGTGGGAGATCACTTTGAGG
rs12260037 Risk allele sense	GATCCCTCAAAGTGATCTGCCACCTCTGCCTCAAAGTGATCTGCCACCTCTGCCTCAA GTGATCTGCCACCTCTGCG
rs12260037 Risk allele antisense	TCGACGAGAGGTGGGCGGATCACTTTGAGGAGAGGTGGGCGGATCACTTTGAGGAGGA GGTGGGCGGATCACTTTGAGG
rs56087297 Normal allele sense	GATCCATGAGTGTATACGTTGTCTGGCCATGAGTGTATACGTTGTCTGGCCATGAGT GTATACGTTGTCTGGCCG
rs56087297 Normal allele antisense	TCGACGGCCAGACACAACGTATACACTCATGGCCAGACACAACGTATACACTCATGGCCA GACACAACGTATACACTCATG
rs56087297 Risk allele sense	GATCCATGAGTGTATACGTTGTCTGGCCATGAGTGTATACGTTGTCTGGCCATGAGT GTATACGTTGTCTGGCCG
rs56087297 Risk allele antisense	TCGACGGCCAGACACAATGTATACACTCATGGCCAGACACAATGTATACACTCATGGCCAG ACACAATGTATACACTCATG

rs12255372 Normal allele sense	GATCCCCAGGCAAGAATGACCATATTCTGACCAGGCAAGAATGACCATATTCTGACCAGGC AAGAATGACCATATTCTGAG
rs12255372 Normal allele antisense	TCGACTCAGAATATGGTCATTCTTGCCTGGTCAGAATATGGTCATTCTTGCCTGGTCAGAA TATGGTCATTCTTGCCTGGG
rs12255372 Risk allele sense	GATCCCCAGGCAAGAATAACCATATTCTGACCAGGCAAGAATAACCATATTCTGACCAGGC AAGAATAACCATATTCTGAG
rs12255372 Risk allele antisense	TCGATCAGAATATG GTTATTCTTGCTGGTCAGAATATGGTTATTCTTGCCTGGTCAGAATATGGTTATTCTTGCC GATCCTATACTTCTGCACCTTCAGCCTCCTTATACTTCTGCACCTTCAGCCTCCTTATACTT CTGCACCTTCAGCCTCCTG
rs11196211 Normal allele sense	TCGACAGGAGGCTGAAGGTGCAGAAGTATAAGGAGGCTGAAGGTGCAGAAGTATAAGGAG GCTGAAGGTGCAGAAGTATAG
rs11196211 Normal allele antisense	GATCCTATACTTCTGCAACTTCAGCCTCCTTATACTTCTGCAACTTCAGCCTCCTTATACTT CTGCAACTTCAGCCTCCTG
rs11196211 Risk allele sense	TCGACAGGAGGCTGAAGGTGCAGAAGTATAAGGAGGCTGAAGGTGCAGAAGTATAAGGAG GCTGAAGGTGCAGAAGTATAG
rs11196211 Risk allele antisense	GCTGAAGGTGCAGAAGTATAG

Table 2. Sequence of oligonucleotides for EMSA

Oligonucleotide names	Primer sequences
AP-1 consensus EMSA probe sense	5'-CGCTTGAT <u>GACTCAGC</u> -3'
AP-1 consensus EMSA probe antisense	5'-TTCCGGCT <u>GAGTCATCAAG</u> -3'
SNP rs7074440 normal EMSA probe sense	5'-TTACAGGTGTGAGTCACCG-3'
SNP rs7074440 normal EMSA probe antisense	5'-CAGGAGCGGTGACTCACAC-3'
SNP rs7074440 risk EMSA probe sense	5'-TTACAGGTGTGAATCACCG -3'
SNP rs7074440 risk EMSA probe antisense	5'-CAGGAGCGGTGATTCACAC-3'

(binding of AP-1 site underlined)