

ゲノムワイド関連解析と エピゲノム解析による 高血圧の機序解明

竹内史比古、清水華、加藤規弘
国立国際医療研究センター研究所

第39回日本高血圧学会総会 SY9-1

2016.10.01 @仙台国際センター

パワーポイント <http://fumihiko.takeuchi.name>

日本高血圧学会 COI開示

国立国際医療研究センター 竹内史比古

演題発表に関連し、開示すべきCOI関係にある企業
などはありません。

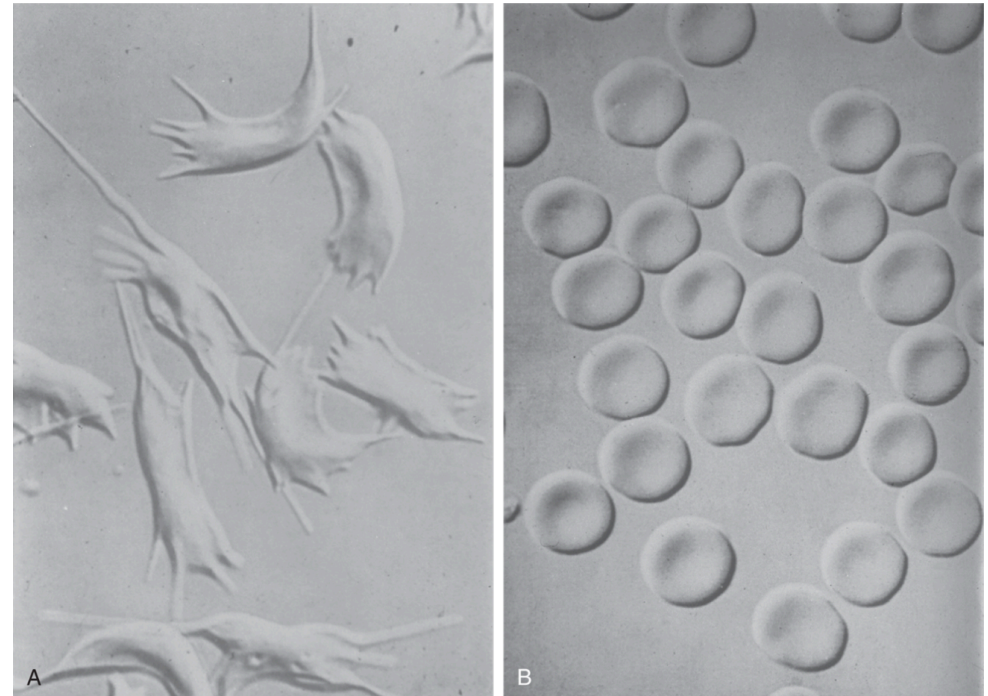
疾患感受性遺伝子を見つけて病気を解明する

- 疾患感受性遺伝子とは
 - DNA変異により、その病気の罹り易さ(感受性)が変わる遺伝子
- 見つける意義
 - 病気の仕組みの解明
 - 創薬ターゲットの探索
 - 個人の発症予測・至適治療法の選択(個別化医療)
- そもそも存在するか？
 - 疾患感受性の素因は、遺伝と環境
 - 家族集積性から遺伝が占める割合(遺伝率)が分かる
 - 高血圧 0.3
 - 身長 0.7
- 疾患感受性遺伝子をごっそり見つけよう→ゲノムワイド関連解析



鎌状赤血球貧血の感受性遺伝子

- *HBB*, ヘモグロビン β 遺伝子
- 一塩基多型 (SNP)
 - ある一塩基に個人差がある
 - c.20A>T (p.Glu7Val) rs334
- 遺伝型は3種類
 - AA
 - 多くの人
 - TT
 - ヘモグロビンが凝集
 - 赤血球が鎌状になる
 - 赤血球が壊され貧血に
 - AT
 - 貧血ない
 - マラリア原虫に感染しにくい





1. ゲノムワイド関連解析

- 方法
- 成功と限界

2. エピゲノム解析

- 組織・機能から
- 機序まで

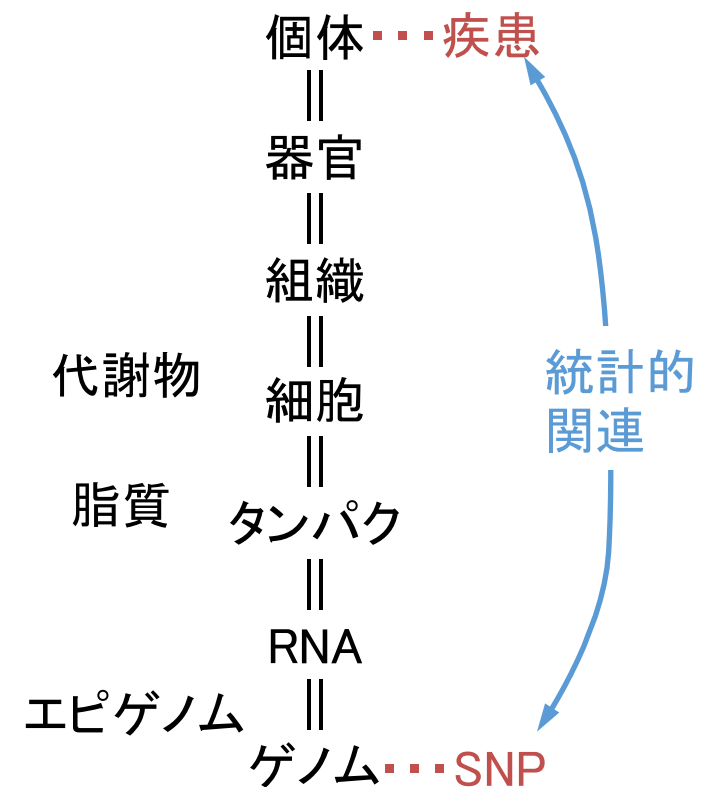
関連解析のコンセプト

- 疾患とSNP

- 生体階層構造の両端に離れている
- 統計的関連(相関)が、ヒトでの因果関係を示唆する
- 中間は、ブラックボックスとしてよい

- 関連解析

- 疾患と関連するSNPを見つける
- 関連SNPの位置にある遺伝子(転写因子結合領域も含む)が、疾患感受性遺伝子のはず
- ゲノムワイドに関連SNPsを探索するのが、ゲノムワイド関連解析(GWAS)



ゲノムワイド関連解析 (GWAS)

- 目標

- ゲノムワイドに、ありふれた(頻度 $\geq 1\%$) SNPs全てについて疾患との関連を検定する

- 計測

- 代表的なSNPsをマイクロアレイで測定し、残りのSNPsの情報は推測

- 統計解析

- SNPsは 6×10^6 個あるが、染色体上で近傍のものは相関している(連鎖不平衡)ので、独立なものは正味 10^6 個
- 約 10^6 回の多重検定を行うので、擬陽性を抑えるために、**有意水準**を $0.05/10^6 = 5 \times 10^{-8}$ と厳しくしないとイケない
- 検出力を上げるためには、罹患者・健常者を**数千人～数万人**調べる必要がある

ゲノムワイド
に網羅的に
調べる



有意水準を
厳しくする

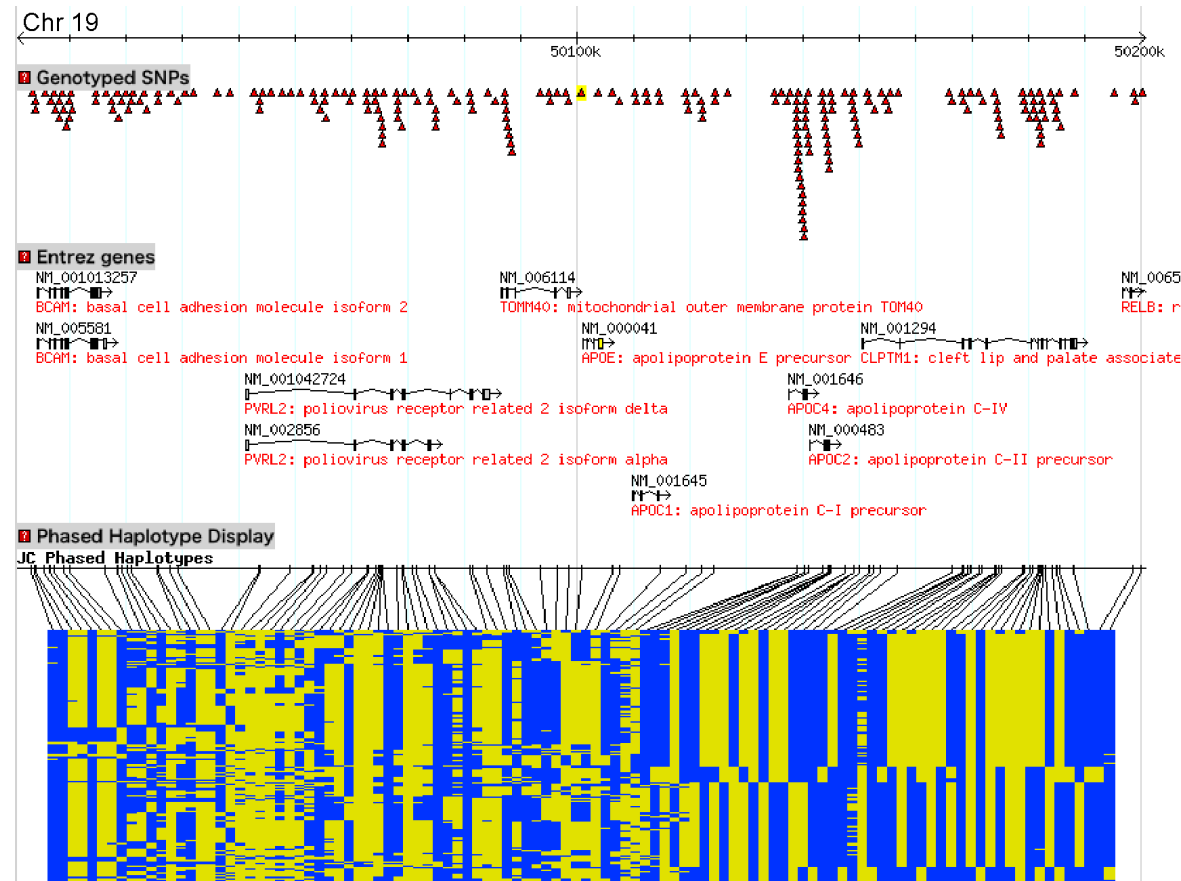


多数のサン
プルが必要

→ 多施設のGWAS研究のメタ解析が現在は主流

SNPs同士の相関(連鎖不平衡)

染色体19番の200kbの領域



日本人45人
(染色体90本)

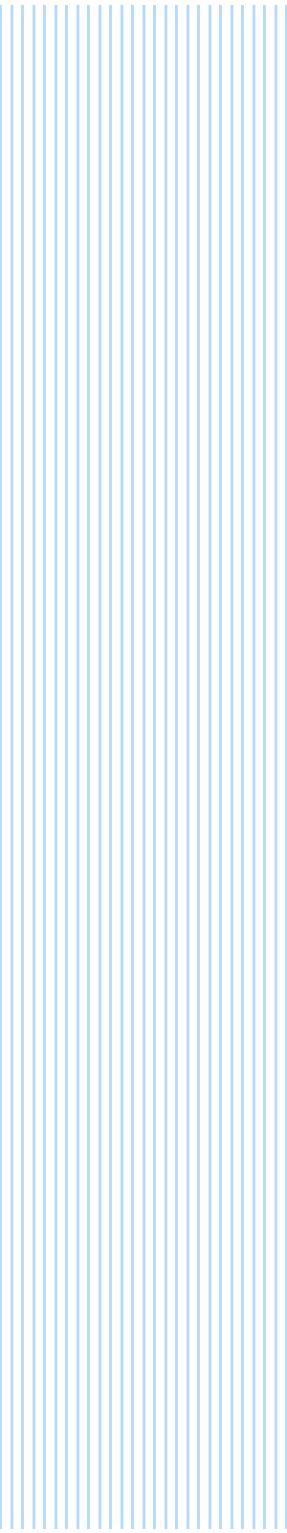
108 SNPs

GWASに必要なサンプルサイズ

- 有意水準 5×10^{-8} のもとで、検出力を 80% にするには、 $40/R^2$ 人必要
 - R^2 (決定係数): 疾患・形質 y の分散のうち、SNP 遺伝子型 x で説明される割合。これは相関係数の二乗。

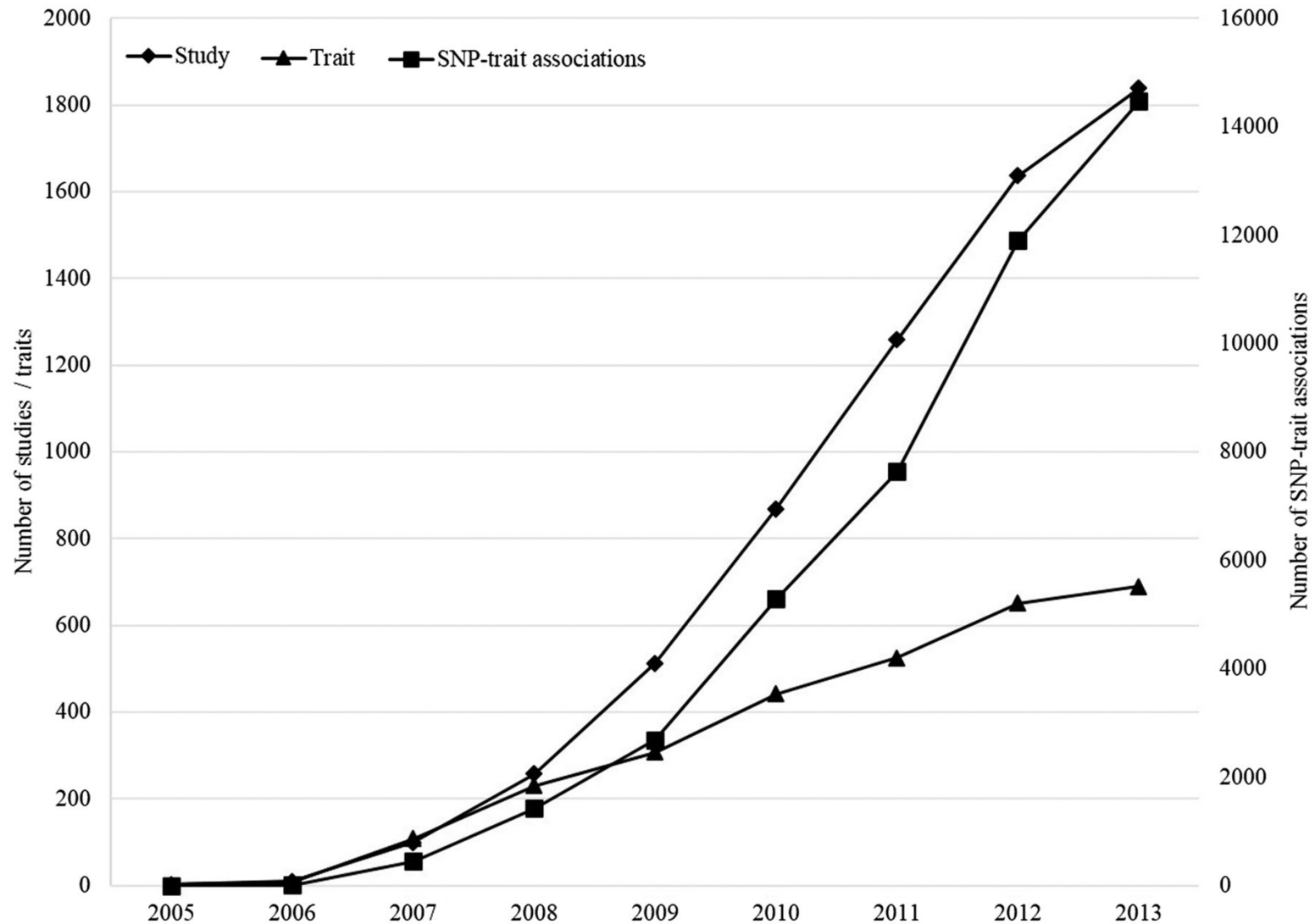
SNPのSBPへの効果 [mm Hg]	R^2 [%]	必要なサンプルサイズ [人数]
1	0.13	32,000
0.5	0.06	64,000
0.25	0.03	128,000

- $F_{1,N-2}$ 分布の非心度パラメータが40になればよい
- SNPの効果が β 、アレルの頻度が p のとき、 $R^2 = 2 p (1-p) \beta^2$
- アレル頻度0.5、SBPの標準偏差を20 mmHgとした

- 
1. ゲノムワイド関連解析
 - 方法
 - **成功と限界**
 2. エピゲノム解析
 - 組織・機能から
 - 機序まで

GWASの成功

<http://www.ebi.ac.uk/gwas/>



2007年頃から普及

Welter et al. NAR (2014) 42:D1001

GWASの成功

<http://www.ebi.ac.uk/gwas/>



The interactive GWAS diagram is a visualization of all SNP-trait associations with $P < 5 \times 10^{-8}$, mapped to the SNP's cytogenetic band.

Welter et al. NAR (2014) 42:D1001

(高) 血圧の大規模GWAS

Study	Publication	年	スクリーニング症例数 [万人]				追試症例数 [万人]	ゲノムワイド有意なSNPsの数	新規SNPsの数
			欧米	東アジア	南アジア	アフリカ			
WTCCC	Nature 447:661	2007	0.5					0	0
Global BPgen	Nat Genet 41:666	2009	3				11	8	8
CHARGE	Nat Genet 41:677	2009	3				3	8	8
AGEN-BP	Nat Genet 43:531	2011		2			3	10	5
ICBP	Nature 478:103	2011	7				13	29	16
COGENT	Am J Hum Genet 93:545	2013				3	10	5	3
iGEN-BP	Nat Genet 47:1282	2015	4	3	3		22	35	12
GHD Exome+, ExomeBP, GoT2D	Nat Genet 電子版	2016	17		3		16	51	30
CHARGE+ Exome	Nat Genet 電子版	2016	12			2	18	70	31
Cardio-Metabochip	Nat Genet 電子版	2016	20				14	66	17
								122	

まとめ: GWASの成功と限界

• 成功

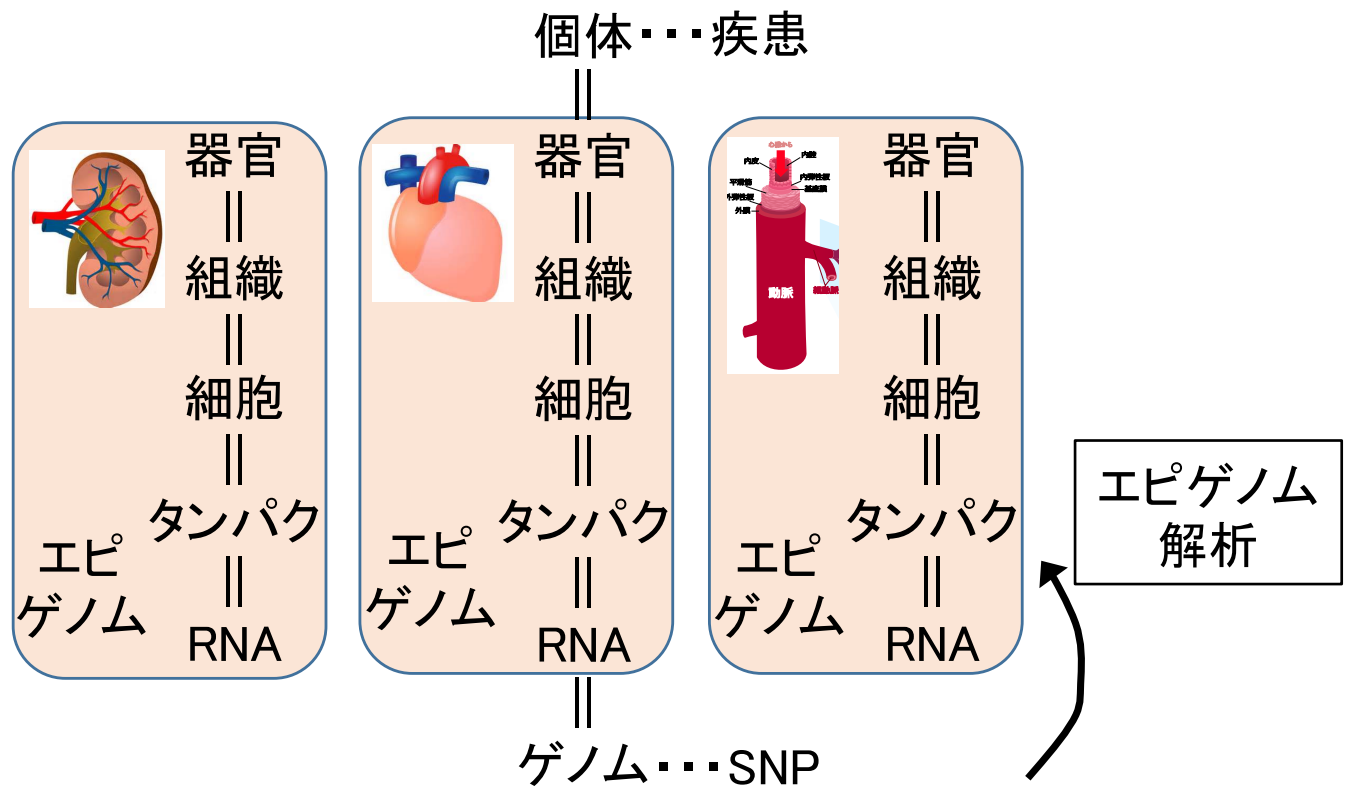
- ゲノムワイド関連解析 (GWAS) では、多数の罹患者と健常者についてSNPsをゲノム全体に渡って測定し、両グループで有意に頻度が異なるSNPsを探索する。
- これまでに数百の疾患や形質についてGWASが行われ、万以上のSNPsとの関連が同定された。
- (高)血圧については、20万人にのぼるGWASから、122 SNPsが同定された。

• 限界

- 互いに連鎖不平衡にある複数の疾患関連SNPsが同定されることが多い。その染色体領域に、複数の遺伝子があるときは、どれが疾患感受性遺伝子か絞り込めない。
- 疾患感受性遺伝子が、どの組織？ どのような機能？ で働くかの解明は、post-GWAS解析が必要。

GWASで説明
できたこと

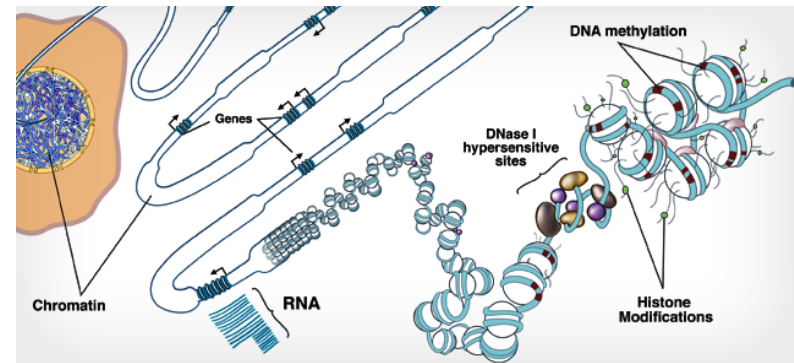
Post-GWASで解析
どの組織？ どのような機能？



エピゲノム解析のアプローチ

- 組織差を調べる

- NIH Roadmap Epigenome Consortium [Nature (2015) 518:317]
- 111もの組織・細胞株 (組織毎のサンプル数は小さい)
- 様々なエピゲノム測定
 - ヒストン修飾、DNAメチル化、DNAアクセス、RNA発現



<http://www.roadmapepigenomics.org/>

- 機能変化を調べる

- エピゲノムワイド関連解析 (EWAS)
 - 数百～数千の症例を調べる。GWASと似た発想
 - EWASでいま現実的なのは、DNAメチル化のアレイによる測定
 - 末梢血を用いることが多い (他の組織の代用とする)
- 細胞・動物実験による機能解析

→ エピゲノム解析により、GWASで同定された疾患感受性遺伝子が、どのような組織・機能で働くか解明されつつある



1. ゲノムワイド関連解析

- 方法
- 成功と限界

2. エピゲノム解析

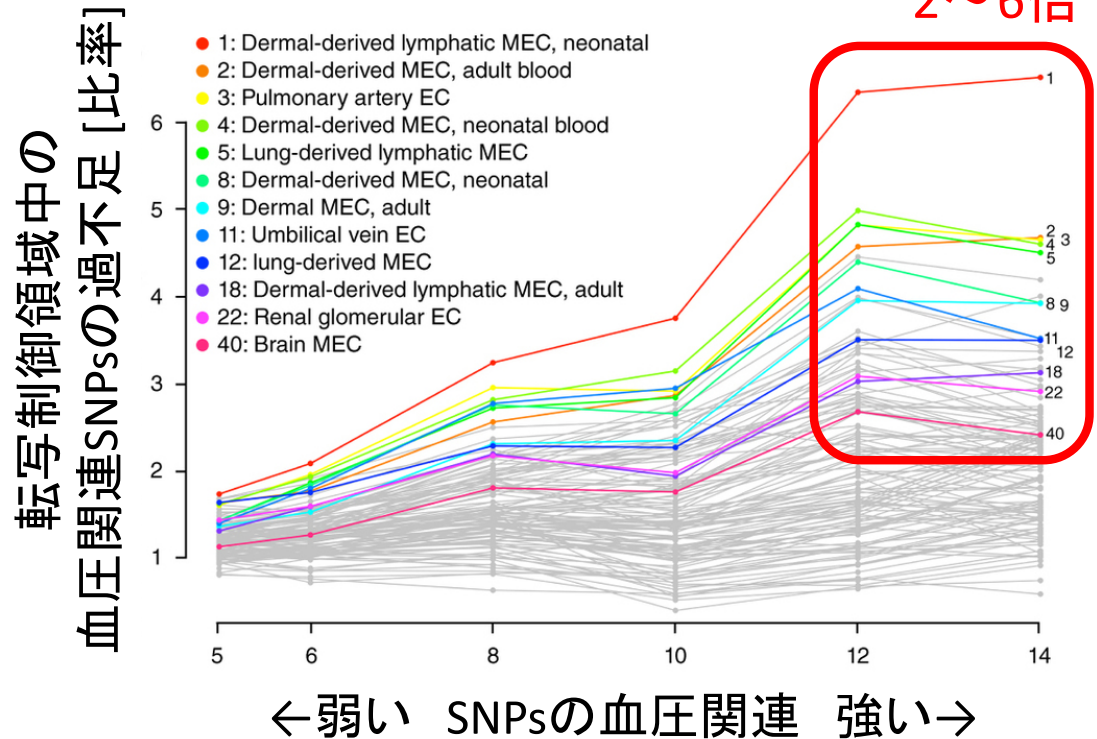
- 組織・機能から
- 機序まで

血圧関連SNPsがどの組織の転写制御領域にあるか

- 123の組織・細胞株のうちどれで、転写制御領域 (DNase I高感受性領域) に血圧関連SNPsが過剰にあるか
 - 血管の組織・細胞株
 - 内皮細胞
- SNPsがこれらの組織で「効いている」

カラー折れ線: 内皮細胞

2~6倍



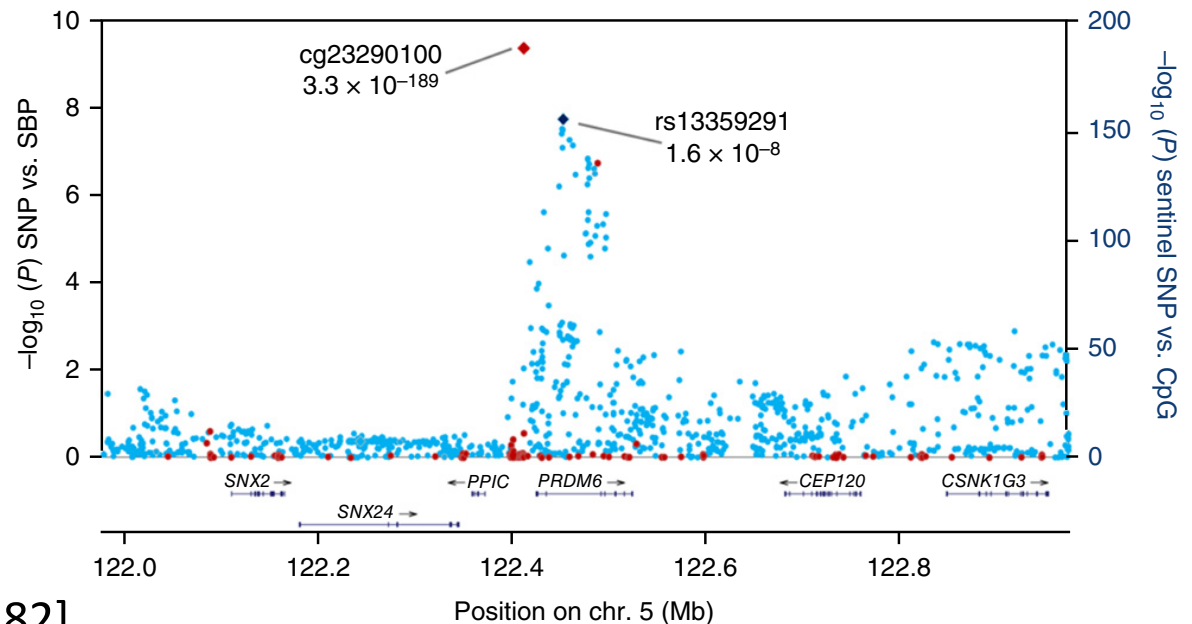
[Ehret et al. (2016)

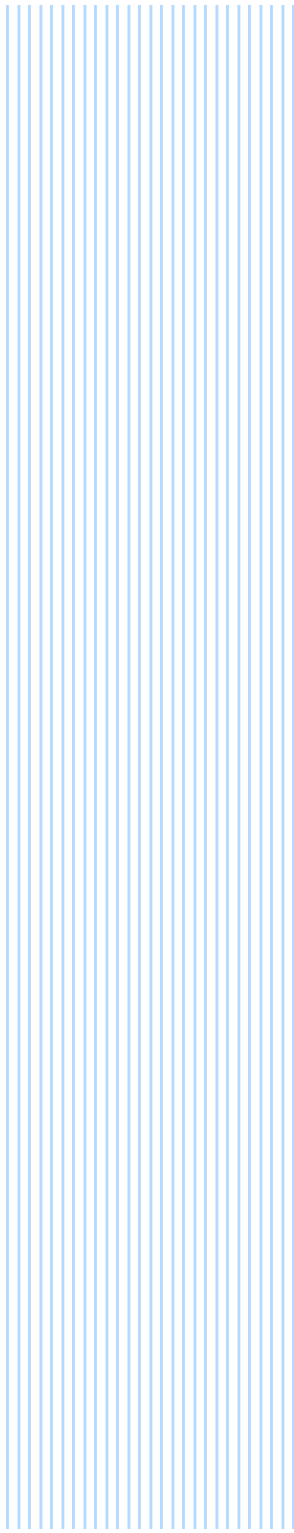
Nat Genet doi:10.1038/ng.3667]

血圧関連SNPsは、近傍のDNAメチル化と関連する傾向がある

- 血圧関連SNPsの周辺1MbのCpGサイトで、メチル化度とSNP遺伝子型が関連するものを探索
- SNP近傍(1kb~244kb)のCpGサイトが関連
- 全体的にみると近傍の関連CpGサイトが多め(ゲノム平均の2倍)
- 「SNP→DNAメチル化→血圧」という経路の機能を示唆

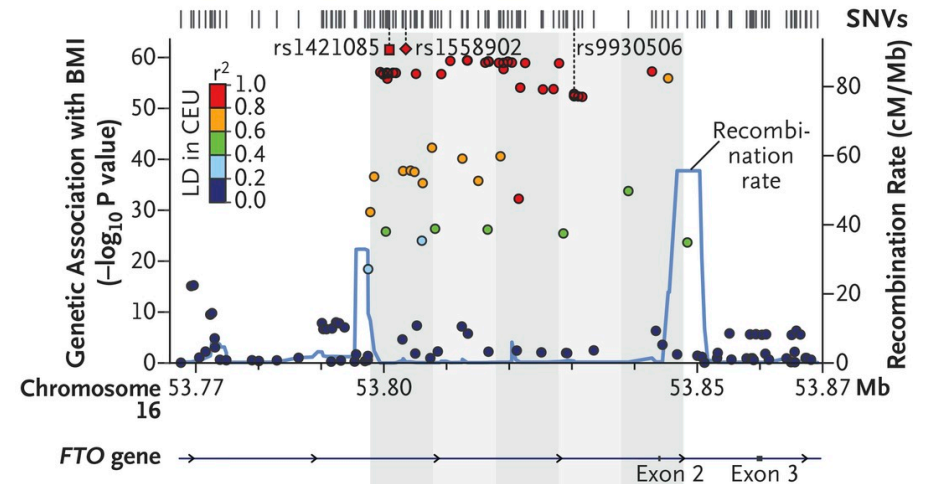
SNP-血圧形質との遺伝的関連 (●)
SNP-CpG siteとの関連 (●)



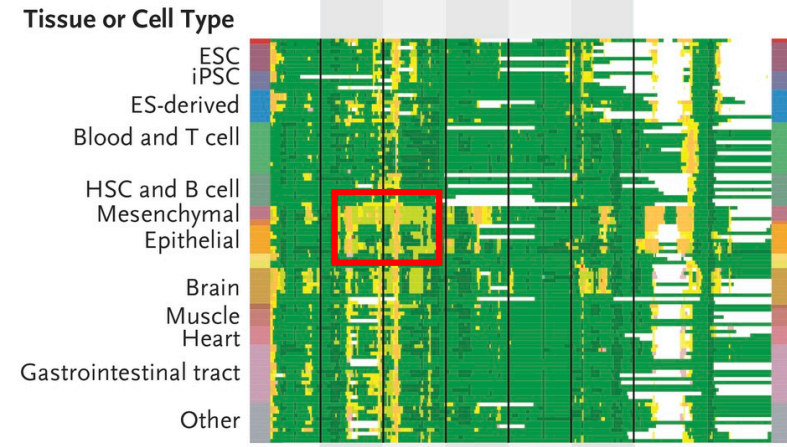
- 
1. ゲノムワイド関連解析
 - 方法
 - 成功と限界
 2. エピゲノム解析
 - 組織・機能から
 - 機序まで

FTO遺伝子のイントロンに位置する肥満関連SNPsのエピゲノム解析

1. 肥満のGWASで、関連SNPsがFTO遺伝子のイントロンに見つかった
2. [組織] この領域がエンハンサーになっているのは、間葉幹細胞・脂肪細胞
3. [機能] SNP rs1421085は転写因子結合部位にある。肥満型アリルで、ARIDモチーフが壊れている
 - 下流のIRX3, IRX5遺伝子を制御
 - 転写因子ARID5Bが原因
4. IRX3, IRX5は脂質貯蔵、ミトコンドリア代謝に関連
5. KO生物で、肥満・やせを確認



3



[Claussnitzer et al. (2015) NEJM 895:907]

FTO遺伝子のイントロンに位置する肥満関連SNPsのエピゲノム解析

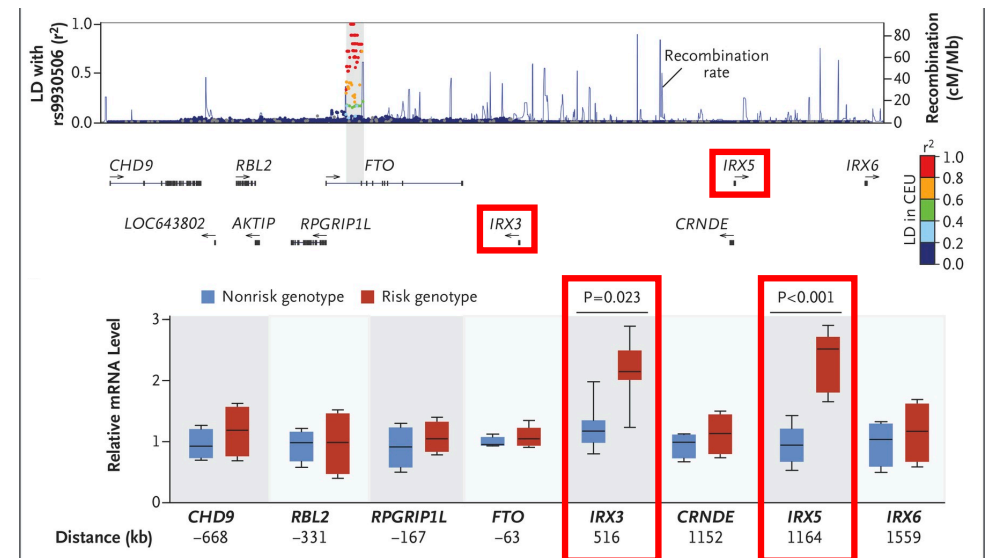
1. 肥満のGWASで、関連SNPsがFTO遺伝子のイントロンに見つかった
2. [組織] この領域がエンハンサーになっているのは、間葉幹細胞・脂肪細胞
3. [機能] SNP rs1421085は転写因子結合部位にある。肥満型アリルで、ARIDモチーフが壊れている

- 下流のIRX3, IRX5 遺伝子を制御
- 転写因子ARID5Bが原因

4. IRX3, IRX5は脂質貯蔵、ミトコンドリア代謝に関連

5. KO生物で、肥満・やせを確認

500kb, 1.2mb 離れている

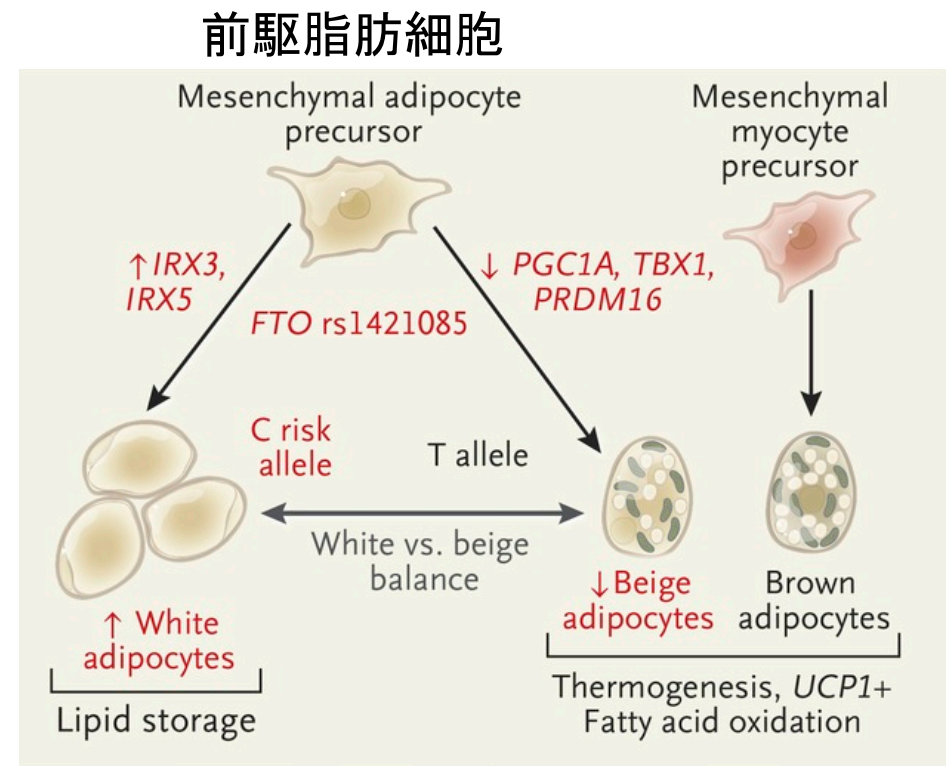


- SGBS脂肪細胞(前駆細胞)で、肥満の遺伝型 vs やせの遺伝型で、IRX3, IRX5が発現変化
- 分化した脂肪細胞では発現変化せず
- FTO遺伝子は関係なかった!

[Claussnitzer et al. (2015) NEJM 895:907]

FTO遺伝子のイントロンに位置する肥満関連SNPsのエピゲノム解析

1. 肥満のGWASで、関連SNPsがFTO遺伝子のイントロンに見つかった
2. [組織] この領域がエンハンサーになっているのは、間葉幹細胞・脂肪細胞
3. [機能] SNP rs1421085は転写因子結合部位にある。肥満型アリルで、ARIDモチーフが壊れている
 - 下流のIRX3, IRX5遺伝子を制御
 - 転写因子ARID5Bが原因
4. IRX3, IRX5は脂質貯蔵、ミトコンドリア代謝に関連
5. KO生物で、肥満・やせを確認



まとめ: post-GWASとしてのエピゲノム解析

エピゲノム解析により、GWASで同定された疾患感受性遺伝子が、どのような組織・機能で働くか解明されつつある。

- 血圧関連SNPsは、血管・内皮細胞などの**組織**で、転写制御領域中に位置していた。
- 血圧関連SNPsは、同時にDNAメチル化とも関連を示しており、DNAメチル化が、SNPと血圧の個人差をつなぐ**分子機能**の一部を成すことが示唆された。
- *FTO*イントロンに位置する肥満関連SNPは、**前駆脂肪細胞**で、**転写因子結合部位を壊す**ことにより、脂質貯蔵・ミトコンドリア代謝に影響し、肥満を起こしていた。

パワーポイント <http://fumihiko.takeuchi.name>