

琉球大学学術リポジトリ

[総説]ヒトゲノム研究の現状と今後の展望

メタデータ	言語: ja 出版者: 琉球医学会 公開日: 2017-08-02 キーワード (Ja): 490 キーワード (En): 作成者: 前田, 士郎, Maeda, Shiro メールアドレス: 所属:
URL	http://hdl.handle.net/20.500.12000/0002016873

ヒトゲノム研究の現状と今後の展望

前田 士郎

琉球大学大学院医学研究科 先進ゲノム検査医学講座

Genetic study for life-style diseases –update and future perspectives–

Shiro Maeda

Department of Advanced Genomic and Laboratory Medicine, Graduate School of Medicine,
University of the Ryukyus

ABSTRACT

After completion of the human genome project, genome-wide association studies (GWAS) have been introduced worldwide, and have identified thousands of confirmed susceptibility loci for many common diseases. In case of type 2 diabetes, over 80 confirmed loci have been identified, and many of them were identified through GWAS in populations of European descent. Some of the European GWAS derived loci have been shown to confer similar susceptibility to type 2 diabetes in non-European populations, including the Japanese. However, integration of all these genetic information can explain only a small percentage (~10%) of the heritability of the disease, and thus most of the genetic factors for type 2 diabetes still remain to be identified, especially in non-European populations. Cumulative evidence suggests that Asians may be more genetically susceptible to type 2 diabetes than populations of European ancestry. Also, there are significant interethnic differences in the risk allele frequency or in effect sizes at several loci, which may affect the power to detect associations in these populations. Therefore, it is considered to be relevant to perform GWAS for type 2 diabetes using non-European populations as well as European populations to uncover the missing heritability of type 2 diabetes. In 2008, two GWAS conducted in Japanese populations have identified *KCNQ1* as a novel susceptibility locus for type 2 diabetes. We further performed an independent GWAS for 459,359 SNPs in 7,541 Japanese individuals (4,470 type 2 diabetes and 3,071 controls), and identified two new loci *UBE2E2* and *C2CD4A-C2CD4B*. Subsequent extended analyses for 2,229,890 imputed SNPs additionally identified *ANK1* locus as a new locus for type 2 diabetes, and additional 3 loci were identified by expanded GWAS examining 26,805 Japanese individuals for 6,209,637 SNPs. Furthermore, in an effort of Asian Genetic Epidemiology Network (AGEN), 8 new loci were identified by a meta-analysis of East Asian GWAS. The associations of *KCNQ1*, *C2CD4A-C2CD4B* and *ANK1* with type 2 diabetes were consistently observed among European populations, underlining the importance of examining non-European populations through GWAS. Further efforts for increasing sample size by scaled-up non-European GWAS or GWAS using isolated populations will help us identify not only ethnicity-specific loci, but also common-susceptibility loci among different ethnic groups, although new approaches other than GWAS, such as whole genome sequencing, will be required to elucidate a heritability of common diseases more precisely. *Ryukyu Med. J.*, 34 (1, 2) 1~8, 2015

Key words: ヒトゲノムプロジェクト, ゲノムワイド関連解析, 2型糖尿病, 1塩基多型

Corresponding Author: 前田士郎. 琉球大学大学院医学研究科 先進ゲノム検査医学講座, 沖縄県中頭郡西原町字上原207番地. Tel : 098-895-3331. E-mail : smaeda@med.u-ryukyu.ac.jp

はじめに

2001年2月にヒトゲノムのドラフトシーケンスが国際コンソーシアムチームと Celera Genomics 社の2つのグループにより「Nature」誌¹⁾と「Science」誌²⁾にそれぞれ発表された。その後、2003年にヒトゲノム配列解読完了宣言³⁾がなされ、2005年には国際ハップマッププロジェクトにより1塩基多型情報およびハプロタイプ地図の整備がおこなわれた^{4,6)}。これらの情報をもとに全ゲノムを網羅した疾患関連遺伝子探索(ゲノムワイド関連解析:GWAS)がスタートし、近い将来個別化医療(personalized medicine)が現実のものになると多くの研究者の期待が高まっていた。2007年にはこのGWASの成果が欧米の研究チームから一気に発表され、その後も精力的にGWASは続けられており、複数の疾患について多くの疾患感受性領域が確立された⁷⁾。糖尿病領域においても2型糖尿病ではすでに80領域程度の感受性遺伝子領域が確立しており、この情報を用いて将来の発症予測に役立てようとの試みも行なわれている。しかしながら、これらの情報すべてを統合しても2型糖尿病の遺伝的要因のわずか1割程度の寄与しか説明できず、臨床の現場に還元するにはほど遠い現状と言わざるを得ない。GWASはありふれた病気の疾患感受性を規定しているのはありふれた個人差(多型)であるという仮説、いわゆる common disease-common variant 仮説に立脚しており、事実GWASで同定された領域は単独での関与の弱い(オッズ比で1.5未満) common variant に限定されている。GWASは common disease の感受性遺伝子研究の重い扉をあけ、大きなブレイクスルーをもたらした事は間違いないが、一方でその限界

もささやかれている。本稿ではGWASを支えるゲノム研究の基本的概念とGWAS後の課題につき述べる。

Common disease-common variant 仮説

メンデルの法則によって説明可能な単一遺伝子病とは異なり、2型糖尿病などのありふれた疾患では、遺伝的要因が存在しても単純なメンデル型遺伝には従っていないように見受けられる。これは単一遺伝子病の遺伝因子(変異)が疾患の発症を規定する決定因子であるのに対し、ありふれた疾患の遺伝因子(多型)は単独効果の弱い(オッズ比で1.1-1.4程度)危険因子にしかならず、しかもこのような遺伝因子が複数存在する上に(多因子遺伝病)個々の症例によって関わる遺伝因子の組み合わせが異なるためと考えられている。多因子遺伝病の遺伝因子はそれ自体重篤な障害に直結するものではなく、したがって選択の影響は殆ど受けない。現代社会では糖尿病などの疾患感受性対立遺伝子として不利益に作用していても、飢餓に曝されていた時代にはむしろ生存に有利に作用していたために多く保存されてきたとする仮説もある(儉約遺伝子仮説、図1)。このような Common disease-common variant 仮説に基づき、単独効果は弱いものの比較的高頻度で存在する複数の遺伝因子を同定する事でありふれた疾患の複雑な遺伝的背景を解明しようとする試みがGWASで行われた。

マーカーとしての1塩基多型

ヒトゲノムの配列のうち0.3%程度に個人間で異なる箇所(バリエーション)が存在している。このうち最も高密度に存在するのが1塩基多型(Single Nucleotide Polymorphism; SNP)であり数百塩基に1ヶ所程度存在する(図2)。SNPは従来の疾患関連

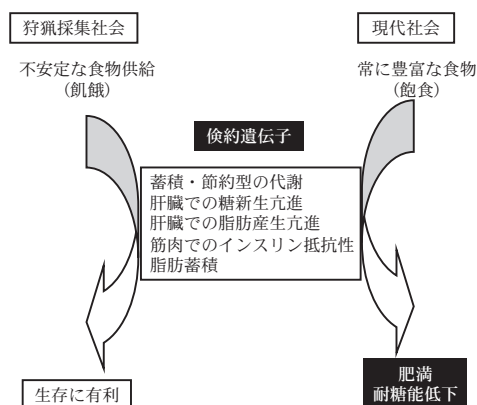


図1 儉約遺伝子仮説 (thrifty genotype hypothesis)

太古の昔飢餓の時代では外来からの摂取糖などを節約する儉約遺伝子型の方が生存に有利であり、選択を受けたために現代社会で数多くの人が保有しているが、現代の飽食の時代ではこの儉約遺伝子型は肥満や糖尿病の発症を促進しているという仮説。

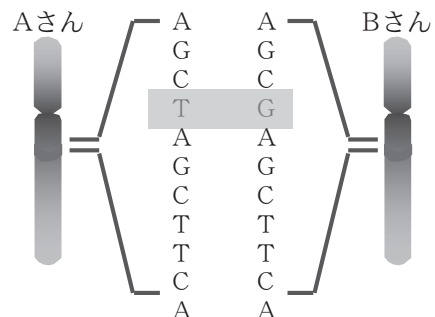


図2 SNP

ヒトゲノム上に存在する1塩基多型(Single Nucleotide Polymorphism[SNP])。数百塩基に1ヶ所認められ、ヒトゲノムには1000万カ所以上存在する。

遺伝子探索のツールであったマイクロサテライトマーカーに比べ高密度に存在する事および通常対立遺伝子は2つ (bi-allelic) であることから大量高速解析に適すると考えられたため、新たな遺伝マーカーとしての役割が期待された。その情報整備はゲノムプロジェクト完了後、まず国際ハップマップ計画 (<http://hapmap.ncbi.nlm.nih.gov/index.html.ja>) で精力的にすすめられた。現在は1000人ゲノムプロジェクト (1000 Genomes: <http://www.1000genomes.org/>)⁸⁾ でさらに整備が継続しており既に1000万カ所以上のSNPs情報が公開されている。

連鎖不平衡マッピング

近接するSNPを詳しく解析してみると、集団における対立遺伝子の存在に強い関連がみられる場合と、全く関連がなく各々独立である場合が存在する。前者を連鎖不平衡、後者を連鎖平衡状態とよび、この状態の形成には減数分裂の際に起こるゲノムの相同組換え、点突然変異の生じた時期、集団の移動あるいは拡散などの因子などが関与しているとされている。結果として一定のゲノム領域内のSNPsは連鎖不平衡状態にあり、この領域を連鎖不平衡ブロックと呼ぶ。連鎖不平衡ブロック内では連鎖不平衡状態にあるSNPsの組み

合わせが複数みとめられ、それぞれの組み合わせを代表するSNPをタグSNPと呼んでいる (図3)。このタグSNPをマーカーとすれば全ゲノム上のSNPsをすべて解析しなくても全ゲノムを網羅することが可能とされている。あるマーカーSNPと疾患との強い関連がみとめられれば、このSNPが存在する連鎖不平衡ブロック内に疾患感受性領域が存在すると考えられる (図4)。このような手法を連鎖不平衡マッピングとよび、さらにこの領域内の多型を詳細に解析することで (ファインマッピング) 原因となる多型の同定が可能とされている。実際には確定した感受性領域とはいえ原因多型を同定する事は容易ではないが、成功すれば疾患感受性メカニズムの解明さらには新規治療法の開発に貢献できるものと考えられる。現在汎用されているSNP解析アレイでは数十万~百万SNPsを同時に解析することが可能である。さらに、この直接解析した遺伝子型情報にデータベースに登録されているゲノム情報を基にしたコンピューター予測 (Imputation) を加える事で実際には解析していない数百万程度のSNPsの遺伝子型を予測する事も可能である。

ゲノムワイド関連解析 (GWAS)

GWASとは疾患に罹患している集団と対照集団と

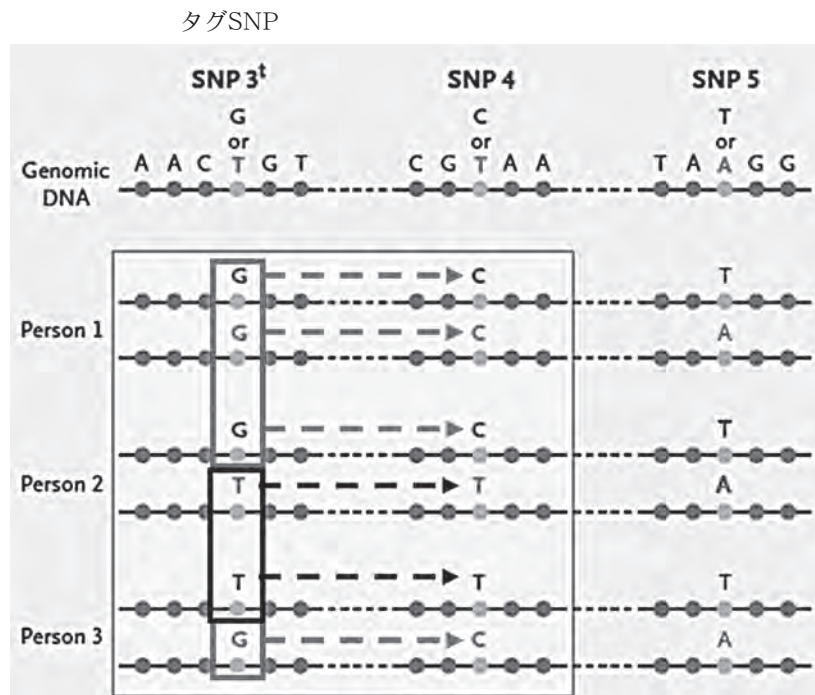


図3 近傍のSNPsと連鎖不平衡

SNP3がGであるとSNP4は必ずCとなっており、SNP3がTであるとSNP4は必ずTとなっている。このような関係にある場合2つのSNPsは絶対連鎖不平衡状態にあるという。完全でなくともある程度この関係が保たれていればこれらSNP集団を代表するSNP (タグSNP) を解析すれば他の同じ集団に属するSNPsの遺伝子型も推測が可能である。

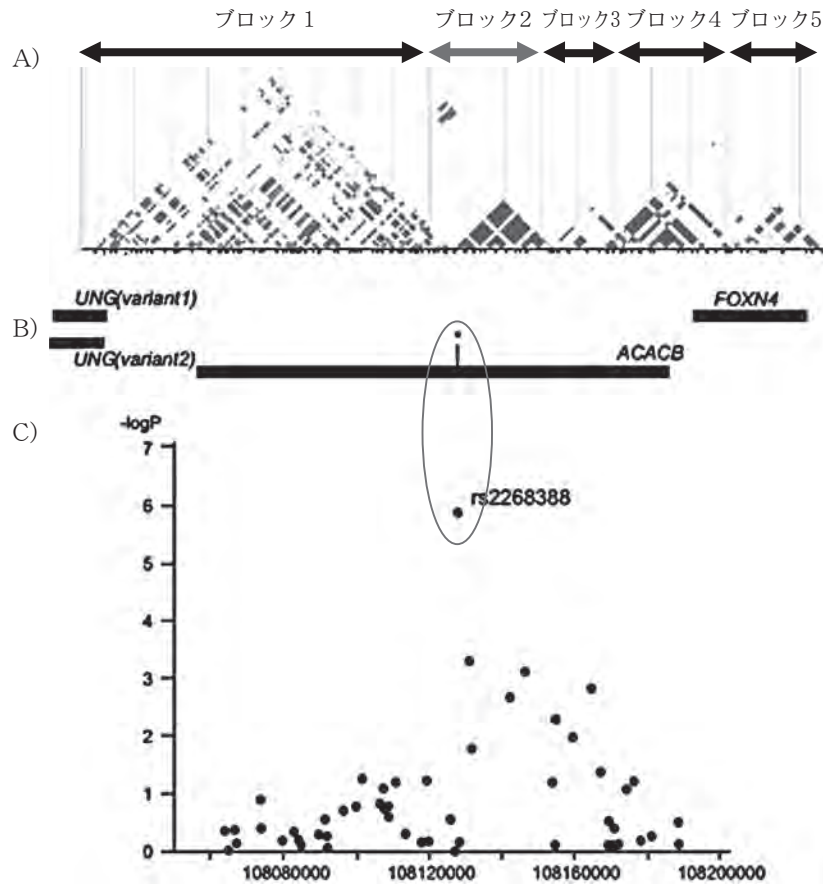


図4 連鎖不平衡マッピング

- A) 連鎖不平衡地図：三角で示されるのがこの領域内の連鎖不平衡ブロック．5つのブロックにわかれている．
 B) この領域に存在している遺伝子
 C) 関連解析の結果：横軸は染色体上の位置，縦軸はP値の常用対数の絶対値．個々のプロットはSNPを表わしており上位にプロットされているほど疾患との関連が強い．Rs2268388が最も疾患と関連が強く，このSNPが存在する連鎖不平衡ブロック（ブロック2）に疾患感受性遺伝子座が存在すると考えられる (Maeda S et al. PLoS Genetics 6: e1000842, 2010).

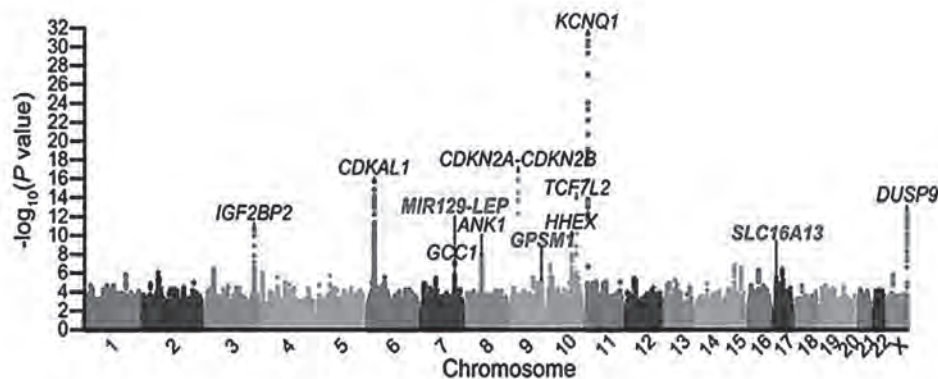


図5 ゲノムワイド関連解析 (GWAS) の解析例

横軸は染色体上の位置，縦軸は関連解析のP値の常用対数の絶対値．この解析では600万カ所以上のSNPsの解析結果が示されている．Manhattan plotと呼ばれており，上方に行くほど関連の強いSNPである (文献14, Hara K et al. Hum Mol Genet 23: 239-, 2014)

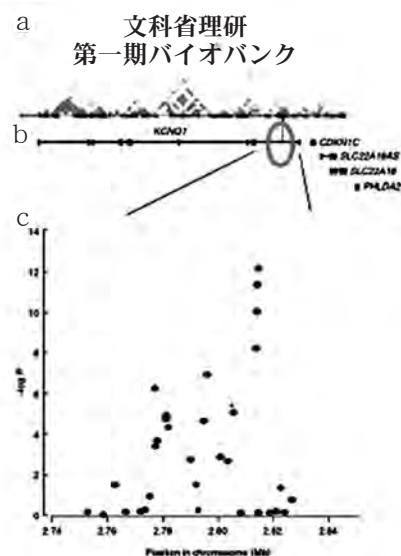
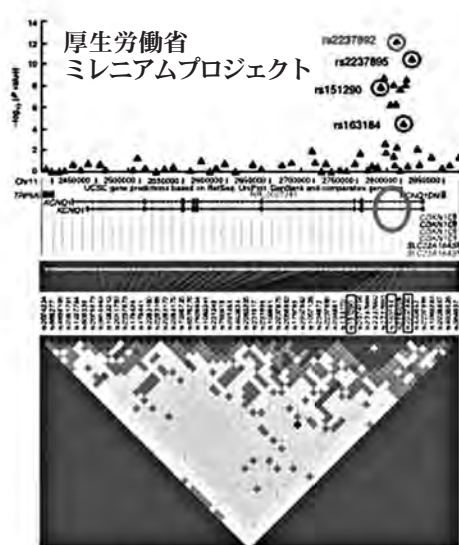
の間でゲノム上の数十万～数百万 SNPs の遺伝子型頻度の違いを検定し (図 5), 疾患の原因となる遺伝子領域探索を全ゲノム領域について行う方法である. 既知の疾患概念にとらわれず統計学的手法によってのみ行うため, 新規の疾患感受性遺伝子領域を同定することができる反面偽陽性が多く検出される可能性がある. そのため, GWAS で確実な感受性領域と認められるためには, 1) 十分な症例数 (数千人以上) で検討すること 2) ゲノムワイド水準の有意差 (p 値 $< 5 \times 10^{-8}$) をみとめること 3) さらに他の独立した集団で再現できること, の 3 つの条件を満たすことが必須とされている⁹⁾.

日本人における 2 型糖尿病 GWAS

我が国における 2 型糖尿病に関するゲノムワイド関連解析はミレニアムゲノムプロジェクトおよびオーダーメイド医療実現化プロジェクトで, それぞれ独立して行われた. 2 つのプロジェクトは同時期に日本人 2 型糖尿病の強力な関連遺伝子領域として 11 番染色体の *KCNQ1* 領域を同定した (図 6)^{10), 11)}. *KCNQ1* 領域は日本人だけでなく, 他の東アジア民族さらには欧米人でも 2 型糖尿病と強く関連していたことから, 人種を超えて重要な 2 型糖尿病疾患感受性領域である事が明らかとなった. *KCNQ1* は心筋に多く発現しているカリウムチャンネルで遺伝性 QT 延長症候群の原因遺伝子として知られていたが糖尿病との関わりは未だ不明である. その後の解析で *KCNQ1* の発現は膵

β 細胞でも確認されておりインスリン分泌調節に関与する事が想定されているが, この領域の責任遺伝子が *KCNQ1* であるのかも含め, 詳細な 2 型糖尿病疾患感受性機構については明らかではない.

これら 2 つの日本人 GWAS は様々な民族で解析する事の重要性を示したが, 一方で検出力 (解析規模) の点で欧米人 GWAS と比較すると劣っていると言わざるを得なかった. そこで我々は単独の GWAS としては最大規模となる日本人 GWAS を行なった¹²⁾. この解析では 2 型糖尿病 4470 人と対照 3071 人について 459,359 SNPs を解析し上位 100 SNPs について別の 2 型糖尿病 2,886 人と対照 3,087 人で再解析を行なった. その結果過去に同定した *KCNQ1* などの領域に加え, あらたに第 3 番染色体の *UBE2E2* 領域および第 15 番染色体の *C2CD4A-C2CD4B* 領域と日本人 2 型糖尿病との関連が認められた. さらに遺伝子型推定 (imputation) により *ANK1* 領域を同定し¹³⁾, これら新規 3 領域についてもその後の大規模な欧米人解析の結果から欧米人 2 型糖尿病との強力な関連が認められている (図 7). その後, さらに日本人 GWAS の規模を拡大し, 3 領域が新たに同定されている (図 5)¹⁴⁾. この 3 領域については欧米人では明らかな関連は認められていないが, *SLC16A13* 領域はその後メキシコ人集団で強力な感受性遺伝子領域であることが示されている. 現時点で様々な民族の GWAS により 2 型糖尿病では約 80 の領域が同定されている (図 8).



(Yasuda et al. Nat Genet 40: 1092-1097, 2008) (Unoki et al. Nat Genet 40: 1098-1102, 2008)

図 6 日本人 GWAS による *KCNQ1* 領域の同定

2 つの独立した日本人研究において, *KCNQ1* のイントロン 1 5 領域と 2 型糖尿病との関連が認められた. 文献 10) 11)

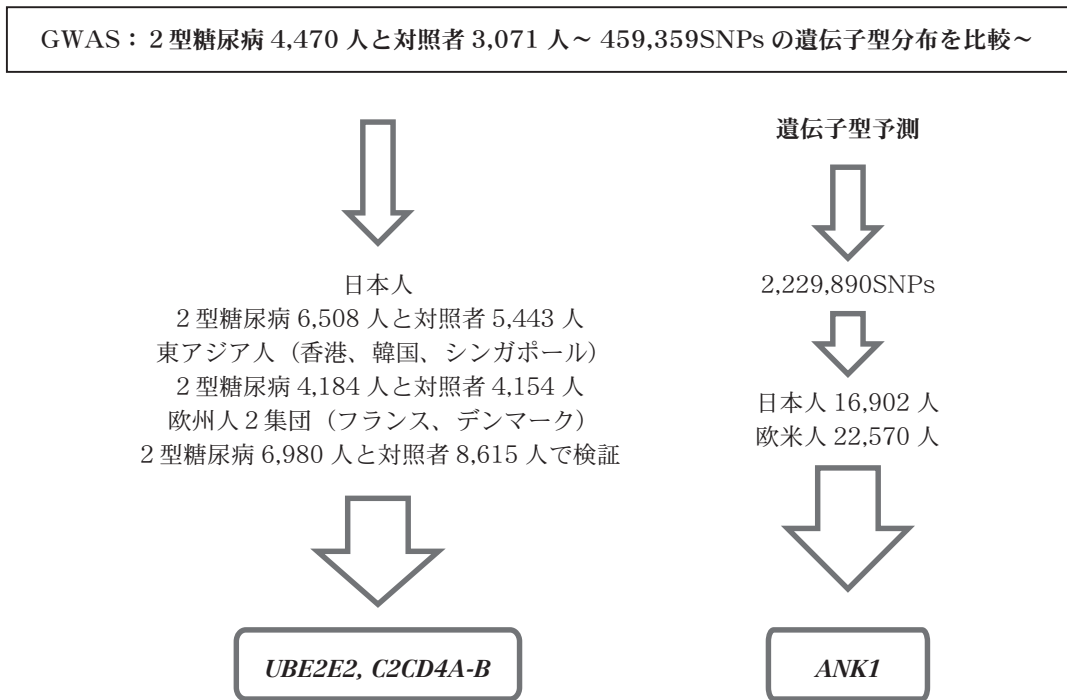


図 7 日本人 GWAS による 3 領域の同定 [文献 1 2), 1 3)]

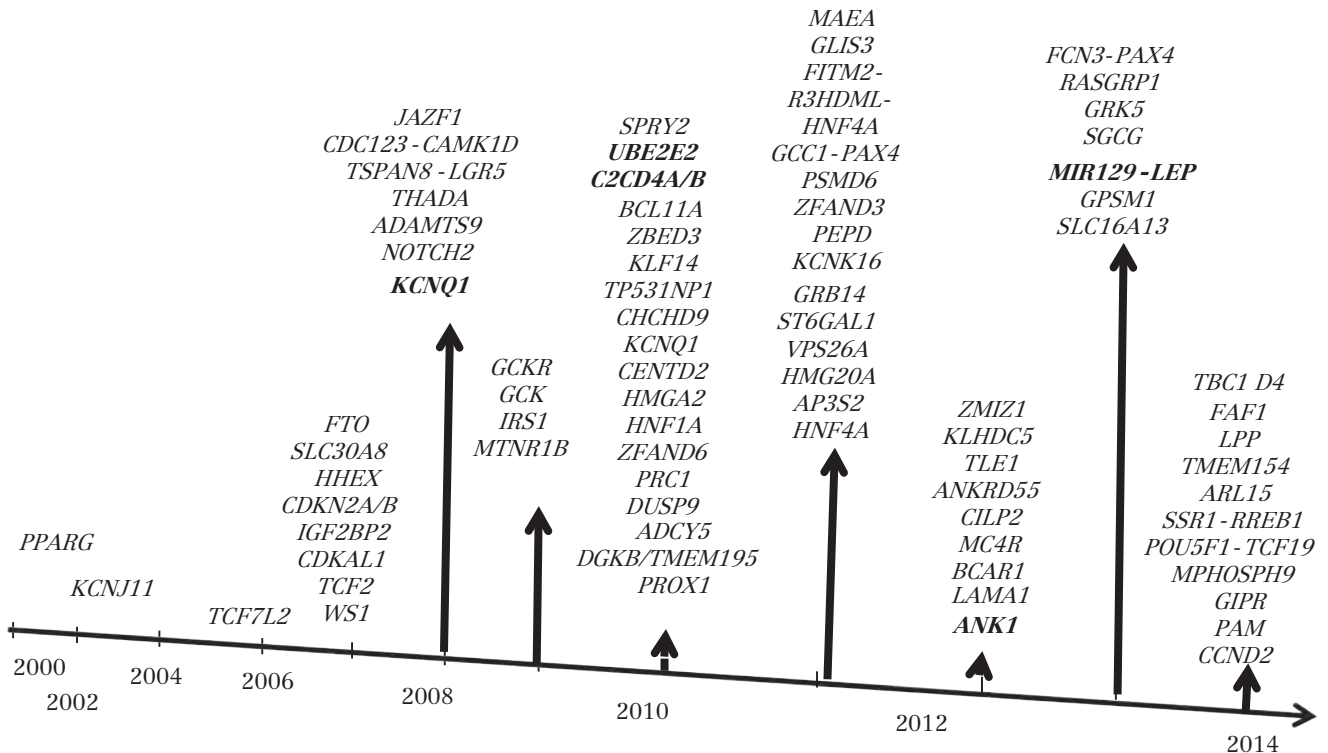


図 8 これまでに同定されている 2 型糖尿病疾患感受性遺伝子領域
 2007 年に GWAS が導入されてから一気に同定が加速している。しかしながら、一部の領域を除き、如何にして 2 型糖尿病疾患感受性に関与するか、その機序に関しては未だ明らかにされていない。
 太字は日本人 2 型糖尿病 GWAS で同定された領域

遺伝学的に独特な集団を用いた解析の有用性

グリーンランドのイヌイット民族を詳細に解析した結果から、2型糖尿病の新たな疾患感受性遺伝子として *TBC1D4* が同定されている。グリーンランド人では、この遺伝子内にナンセンス変化をおこす variant (Arg684Ter, c2050C>T, rs6176969) が 17% で認められ、この variant のホモ接合体では 2 型糖尿病のオッズ比は 10.3 となると報告されている。一方、他の欧米人、東アジア人、アフリカ人ではこの 684Ter variant は殆ど認められない¹⁵⁾。*TBC1D4* はその後、骨格筋のインスリン抵抗性に関与する事が明らかとなり、グリーンランド人のみならず、全世界の 2 型糖尿病患者の新規治療標的となる事が示されている。このように遺伝学的に孤立した集団においては、比較的効果の強い variant が保存されている可能性がある。沖縄県は島嶼県であり、そのなかでも宮古島は本島、八重山とは遺伝学的に異なった集団を形成している事が報告されており¹⁶⁾、今後イヌイット解析と同様の手法により生活習慣病の新規領域同定が期待できると考えられる。

おわりに

当初、GWAS により多くの common disease の感受性領域同定が達成されれば個別化医療が現実のものとなると多くの研究者が考えていた。しかしながら皮肉なことに多くの領域が同定されるに従い、個別化医療実現の困難さをさらに思い知らされる結果となっている。さらに同定された領域がいかんして疾患感受性に関わるかその詳細な機序に関しては殆ど解析が進んでいない。薬剤の副作用予測などではすでに有用な情報も存在してはいるものの、個別化医療時代の到来にはまだかなりの年月と労力が必要なようである。道のりは長いと言わざるを得ないが、今後、特に欧米人以外のより大規模な GWAS の結果から、さらなる感受性遺伝子領域同定がすすみ、次世代シーケンサーを駆使したより高度な解析さらにはエピゲノム情報や腸内細菌叢などのメタゲノム解析情報などの統合解析により着実に進歩を遂げる事で、個別化医療実現が達成される日が来る事を期待したい。

参考文献

1) International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* 409: 860-921,

2001

- 2) Venter JC et al, *Science* 291: 1304-1351, 2001
- 3) International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945, 2004
- 4) The International HapMap Consortium, A haplotype map of the human genome. *Nature* 437: 1299-1320, 2005
- 5) International HapMap Consortium, Genome-wide detection and characterization of positive selection in human populations. *Nature* 449: 913-918, 2007
- 6) International HapMap Consortium, A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449: 851-861, 2007
- 7) Lander ES, Initial impact of the sequencing of the human genome. *Nature*, 470: 187-197, 2011
- 8) The 1000 Genomes Project Consortium, A map of human genome variation from population-scale sequencing. *Nature* 467:1061-1073, 2010
- 9) Imamura M et al, Genetics of type 2 diabetes. GWAS era and future perspectives. *Endocrine J* 58: 723-739, 2011
- 10) Unoki H, Takahashi A, Kawaguchi T et al. SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat Genet* 40: 1098-1102, 2008.
- 11) Yasuda K, Miyake K, Horikawa Y et al.: Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet* 40: 1092-1097, 2008.
- 12) Yamauchi T, Hara K, Maeda S et al.: A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at *UBE2E2* and *C2CD4A-C2CD4B*. *Nat Genet* 42: 864-868, 2010
- 13) Imamura M et al.: A single-nucleotide polymorphism in *ANK1* is associated with susceptibility to type 2 diabetes in Japanese populations. *Hum Mol Genet* 21:3042-3049, 2012
- 14) Hara K et al.: Genome-wide association study identifies three novel loci for type 2 diabetes. *Hum Mol Genet* 23:239-246, 2014
- 15) Moltke I et al.: A common Greenlandic *TBC1D4* variant confers muscle insulin resistance and

type 2 diabetes. *Nature* **512**:190-193, 2014
16) Sato T et al. Genome-wide SNP analysis reveals population structure and demographic history

of the Ryukyu islanders in the southern part of the Japanese archipelago. *Mol Biol Evol* **31**:2929-2940, 2014