

Annual Review 2011

腎 臟

編 集 | 富野康日己 順天堂大学教授
柏原 直樹 川崎医科大学教授
成田 一衛 新潟大学教授

中外医学社

1. GWAS

新潟大学大学院医歯学総合研究科腎膠原病内科学 後藤 眞
同 教授 成田一衛

key words GWAS, SNP, common variant, rare variant, CNV

動 向

単一遺伝子異常によるメンデル遺伝疾患に対して、高血圧や糖尿病などの多因子疾患には、複数の遺伝因子が関与していると考えられている。ゲノムプロジェクトによりヒトゲノムの解読が終了し、個々人のSingle Nucleotide Polymorphism (SNPs) が明らかにされ、ありふれた疾患の原因を common variant に求めて全ゲノム領域にわたる関連解析 Genome-Wide Association Study (GWAS) が行われるようになった。SNP タイピング技術の進歩と HapMap プロジェクトにより連鎖不平衡ブロックやタグ SNP が明らかにされ、2007 年以降、GWAS に関する報告は飛躍的に増加した。しかし、GWAS による知見が蓄積される一方で、検出された疾患感受性リスクアレルの効果は不十分であり、疾患遺伝率の一部しか説明できないことが明らかとなった (missing heritability)。このことから、アレル頻度が低く、GWAS では検出できないが、効果が大きい rare variant の役割が大きいと推測されている。また遺伝子間相互作用、遺伝子・環境相互作用、エピジェネティクスなども今後明らかにされていくと思われる。シーケンス技術のめざましい発展により個人の全ゲノム情報をシーケンスで決定することも可能な時代となり、1000 人ゲノムプロ

ジェクトとともに疾患に関係するあらゆる variation がカタログ化されることも近いと予想される。

A. GWAS のデザインと解析

GWAS を行ううえで疾患の表現型を正確に把握することは非常に重要である。また遺伝的に異なる集団が混在しているとアレル頻度が異なり、偽陽性が生じてしまうため集団の階層化に注意してサンプルを収集する¹⁾。GWAS に必要なサンプル数は対象とする疾患リスクアレルの頻度、効果により異なる。初期の GWAS では 1000 以下のサンプル数でオッズ比が 2 以上のリスクアレルを同定する例がみられたが、多くの多因子疾患のオッズ比は 1.1 から 1.4 であり、検出力 (一般的には 0.8 以上) を備えた GWAS を計画するためには、リスクアレルの頻度により 8,000 人から 20,000 人のサンプル数が必要となる²⁾。統計学的に全ゲノム領域から疾患の原因となる SNPs を同定することは慎重にならなければいけない。全ゲノム領域で多数の SNPs を比較検定する際、多重比較による偽陽性を避けるためには解析結果を補正する必要がある。一般的に Bonferroni 法による補正からゲノムワイド有意水準を $P=10^{-7} \sim 10^{-8}$ とする報告が多いが、保守的な検定のため検出力は低

下する。SNPsの頻度やSNPs間の相互作用などにより有意水準は変化すると予想される³⁾。

GWASで一次スクリーニングを行った後に、統計学的に有意なSNPsについては別の集団で再現性を確認する必要がある。再現性が得られた領域についてはfine-mappingを行い、common SNPsからrare SNPsまで疾患の原因となるvariationを機能的に解析する。集団は多様な集団を対象にすることが望ましい。集団によっては疾患の頻度が異なり、リスクアレルの頻度や効果も異なる可能性があり、また、連鎖不平衡の構造によってはリスクアレルの検出が容易になる⁴⁾。近年、タイピングされていない遺伝子型を、HapMap referenceデータから推測するimputation法が行われている⁵⁾。imputation法により検出力の上昇、メタ解析やfine-mappingでの原因SNPの同定に効果を発揮している。

B. GWASにより明らかにされた知見

2007年以降、GWASが盛んに行われるようになり、2010年3月末までに148の疾患に対して779の研究論文が報告された (<http://www.genome.gov/GWASudies/>)。GWASにより疾患の病態生理の理解が進んだことは事実であり、関連すると推定されていた遺伝子内のSNPsが、GWASにより実際に検出された。LDL-C、HDL-C、中性脂肪の値に関するGWASでは脂質代謝に関連する12の遺伝子座が検出され、LDL-Cの値に関しては、HMGCoA還元酵素のSNPがリスクアレルとして報告された^{6,7)}。一方、GWASにより、新しい病態生理が明らかになった例も認められる。加齢黄斑変性症における補体制御因子H⁸⁻¹⁰⁾、乳癌でのFGFR2¹¹⁾、2型糖尿病のCDKN2AやCDKN2B^{12,13)}、などであり、GWASが新しい研究領域を開拓した最も実りある例と言える。また類似した臨床所見を有する疾患群では、GWASにより共通した領域で関連が

認められた。クローン病、乾癬、強直性脊椎炎では、IL23受容体で共通のリスクアレルを有し¹⁴⁾、同様に関節リウマチとSLEの両者ではSTAT4のSNPsで関連が認められた¹⁵⁾。今まで関係がないと思われていた疾患群で共通した遺伝子座を有する例も判明した。CDKALIにおいて2型糖尿病とクローン病は共通したリスクアレルを有し¹⁶⁾、またTCF2において2型糖尿病と前立腺癌¹⁷⁾、TCF7L2において2型糖尿病と大腸癌が共通して関連が認められている¹⁸⁾。

一方で関連解析で検出された領域に遺伝子群が認められない疾患も多い¹⁹⁾。これらは蛋白をコードしていない領域であり、遺伝子発現を制御している領域の可能性もある。近年はマイクロアレイを用いた遺伝子発現プロファイルとQTL解析を組み合わせた発現QTL解析も行われ²⁰⁾、関連解析で検出されたリスクアレルにより制御される遺伝子発現機序が研究されている。

C. 腎臓領域でのGWAS

2009年に腎機能（血清クレアチニン、シスタチンC、MDRD式による推定GFR）、慢性腎臓病（CKD）について19,877名を対象にしたGWASが報告された。eGFRやCKDとUMODに最も強い関連が認められた²¹⁾。UMODはヘンレ係蹄の太い上行脚で発現するTamm-Horsfall蛋白をコードするが、CKDの病態での関与は明らかでなかった。2010年にはKottgenらが約20のGWASをメタ解析し、67,093名を対象として腎機能あるいはCKDと有意に関連する新規の23アレルを検出し、さらに22,982名による再現性の解析で13のアレルが関連することを報告した²²⁾。これらのアレルを含む遺伝子は腎発生、糸球体スリット膜、糸球体上皮細胞、血管新生、溶質トランスポーターに関連し、機能的に興味深い。Chambersらも、23,812名を対象に腎機能についてGWASを行い、血清クレアチニンと2p12-p13, 6q26,

17q23, 19q13におけるアレルが有意に関連し、中でも2p12-13と19q13はeGFR, シスタチンC, CKDとも関連が認められたと報告した²³⁾. 2p12-13と19q13はKottgenらの報告した遺伝子座と一致し、腎尿細管に特異的に発現して腎代謝に関係する酵素 (*NAT8*) や尿細管でのアミノ酸輸送に関係するトランスポーター (*SLC7A9*) が含まれ、今後、CKDと代謝、溶質輸送についての研究が進むと思われる。

糖尿病性腎症については、1型腎症に関するGWAS (820症例と885名のコントロール) で *FRMD3* (細胞骨格蛋白) と *CARS* (tRNA合成酵素) で関連が認められ、DCCT/EDICでの前向きコホートで、リスクアレルを有する症例は有意に高率に腎症を発症した (HR = 1.3)²⁴⁾. 2型腎症については、日本において全ゲノム領域のcSNPを中心とするデータベース (JSNP) で関連解析が行われ、*ELMO1*などで関連が認められているが²⁵⁾、さらに最近、1,312人の2型糖尿病性腎症を対象とし、アセチルCoAカルボキシラーゼ (*ACACB*: 遊離脂肪酸の代謝にかかわる酵素) において欧州の集団とともに関連が認められ、オッズ比は1.6と報告された²⁶⁾.

IgA腎症については、2010年に、欧州からIgA腎症に関する初めてのGWASが発表された²⁷⁾. 家族ベースの関連解析に187症例、症例対照・関連解析に244症例を対象として、約32万SNPsをタイピングし、両サンプルを合わせた解析で6番染色体短腕のMHC領域で $P < 10^{-9}$ を示す結果が得られ、HLA領域にIgA腎症の発症に関与する疾患感受性アレルが存在することが報告された。日本でもJSNPを用いた解析でHLA-DRAに有意な関連を認めており²⁸⁾、HLAとIgA腎症の発症についてさらなる解明が期待される。

D. Missing heritability

GWASにより明らかにされたリスクアレル

(common variant) の遺伝効果は、ごく少数の疾患を除いて一般的に小さい。アルツハイマー病でのAPOE4, 加齢黄斑変性症での補体因子H, 剥離性緑内障におけるLOXL1, などのリスクアレルは2以上のオッズ比を有するが、それ以外の関連解析で検出されたリスクアレルのオッズ比は1.1~1.5にとどまっている²⁹⁻³¹⁾. そしてこれらの効果を総和してもGWASにより同定されたアレル (common variant) は、各々の疾患の遺伝率のごくわずかしが説明できない (加齢黄斑変性症では約50%だが、クローン病では約20%, 2型糖尿病では約6%と推定されている). 最も定量的な形質である身長についても *HMGA2* を含めた40以上のリスクアレルが報告されているが、60~80%とされる遺伝率の5%程度しか説明できない^{29,31)}. さらに効果の小さいリスクアレルがまだ検出されていないか、あるいは、頻度は少ないがより効果の大きいリスクアレルがGWASで検出されていない可能性がある。頻度が低いvariantは、low frequency variant (0.5~5.0%) や rare variant (<0.5%) に分類されるが³²⁾, rare variantもありふれた疾患の病態に関与していることは明らかである。例として、ATP結合輸送膜蛋白であるABCA1遺伝子はコレステロール排出トランスポーターとして機能するが、重度の低HDLコレステロール血症ではまれな非同義変異が多く認められており、common variantに加えてrare variantが病態生理に関与している³³⁾. 現在のHapMapプロジェクトでは主にアレル頻度が5%以上のcommonSNPsを対象としており、頻度が低いSNPsのカバー率は低く、現在のGWASではrare variantを検出できない。GWASで検出されたcommon variantの周辺でリシークエンスを行ったり、形質 (疾患) のなかでも極端な表現型を示す症例についてシークエンス解析を行うことで効果の大きいrare variantを検出できる可能性がある³⁴⁾ (図1)。

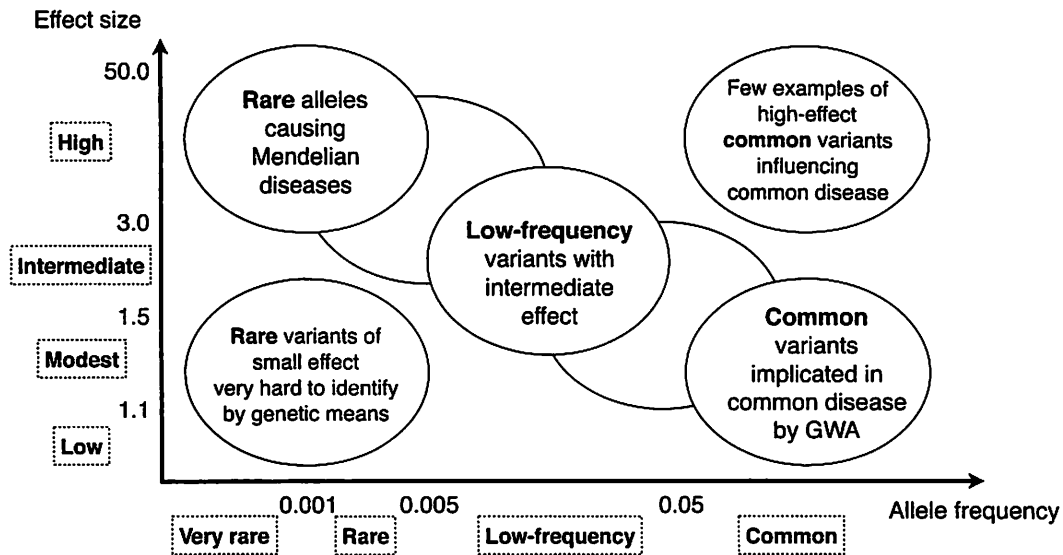


図1 頻度とオッズ比からみたリスクアレル (Nature. 2009; 461; 747³²) より一部改変)

近年, SNPアレイの解析から新たに Copy Number Variation (CNV) などの構造多型が明らかにされた。頻度が1%以上のCNVは Copy Number Polymorphism (CNP) とされる。自閉症を対象した解析では, 16p11.2の領域で約593kbの de novo 欠失と重複が1.5%の症例で観察され³⁵⁾, 統合失調症などでも rare CNV (欠失)が報告され, いずれもオッズ比は高い³⁶⁻³⁸⁾。一方, クローン病や乾癬ではサイズが小さい common CNPとの関連が報告され, オッズ比は1.5と小さい^{39,40)}。CNVの解析はまだ十分ではないが, 今後, プラットフォームの改良でCNVの検出感度が改良され, rare CNVと common CNPの遺伝率に与える影響も明らかにされると思われる。

GWASにより明らかにされたリスクアレルの危険度には, 遺伝子・遺伝子の相互作用や遺伝子・環境の相互作用は考慮されていない。rare variantやCNVのみが missing heritabilityを説明し得るのではなく, 複合的な相互作用が明らかになれば common variantの危険度はさらに大きくなる可能性がある⁴¹⁻⁴³⁾。

E. 今後のゲノム解析とGWAS

GWASにおいて, common variantからCNVや頻度の低いvariantに広げるためには今まで以上に網羅的なゲノム多型情報のカタログが必要となる。近年, シークエンス技術の革新とともに1000人ゲノムプロジェクトが進行している (<http://www.1000genomes.org/page.php>)。このプロジェクトにより頻度が低いSNPsの詳細なマップが作成され, common variantとrare variantのギャップを埋めてありふれた疾患の原因遺伝子の発見を促進すると期待されている。

これまでのGWASは欧米の集団に偏っており, 他の集団における報告は少なかった。特に人類の起源と考えられているアフリカ人はヨーロッパ人やアジア人に比較して遺伝学的に多様であり⁴⁴⁾, 連鎖不平衡の程度は弱い。GWASによるスクリーニングでは強いLDを有する集団が有利であるが, fine mappingではむしろ弱いLDを有する集団の方が原因SNPを同定しやすい可能性がある。また, rare variantはcommon variantに比べて,

世代数が少なく、限定された集団に存在する可能性もある⁴⁾。今後、アフリカ人も含めた多集団でのGWASの知見が重要であると考えられている⁴⁵⁾。

将来は、連鎖解析やGWASによるゲノムマッピングから、すべての多型情報を収集し、表現型との関連を検討する時代へ移行すると考えられる。さらに多数のvariantと表現型および環境因子との関連を検証する新しい統計学的手法も必要となる。エピジェネティクスから疾患表現型への関与も明らかになるであろう。ゲノム医学から疾患の発症のリスクアレルが明らかになれば、その遺伝子を修飾した動物モデルの開発が可能になり、また臨床試験においてもリスクアレルにより層別化を行い、より治療効果の判定が行いやすくなる可能性がある²⁾

文献

- 1) Price AL, Zaitlen NA, Reich D, et al. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet.* 2010; 11: 459-63.
- 2) Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science.* 2008; 322: 881-88.
- 3) Hoggart CJ, Clark TG, De Iorio M, et al. Genome-wide significance for dense SNP and resequencing data. *Genet Epidemiol.* 2008; 32: 179-85.
- 4) Rosenberg NA, Huang L, Jewett EM, et al. Genome-wide association studies in diverse populations. *Nat Rev Genet.* 2010; 11: 356-66.
- 5) Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet.* 2010; 11: 499-511.
- 6) Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* 2008; 40: 161-9.
- 7) Kathiresan S, Melander O, Guiducci C, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet.* 2008; 40: 189-97.
- 8) Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science.* 2005; 308: 385-9.
- 9) Edwards AO, Ritter Rr, Abel KJ, et al. Complement factor H polymorphism and age-related macular degeneration. *Science.* 2005; 308: 421-4.
- 10) Haines JL, Hauser MA, Schmidt S, et al. Complement factor H variant increases the risk of age-related macular degeneration. *Science.* 2005; 308: 419-21.
- 11) Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature.* 2007; 447: 1087-93.
- 12) Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science.* 2007; 316: 1336-41.
- 13) Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science.* 2007; 316: 1341-5.
- 14) Burton PR, Clayton DG, Cardon LR, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet.* 2007; 39: 1329-37.
- 15) Kobayashi S, Ikari K, Kaneko H, et al. Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population. *Arthritis Rheum.* 2008; 58: 1940-6.
- 16) Wolf N, Quaranta M, Prescott NJ, et al. Psoriasis is associated with pleiotropic susceptibility loci identified in type II diabetes and Crohn disease. *J Med Genet.* 2008; 45: 114-16.
- 17) Gudmundsson J, Sulem P, Steinthorsdottir V, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet.* 2007; 39: 977-83.
- 18) Folsom AR, Pankow JS, Peacock JM, et al. Variation in TCF7L2 and increased risk of colon cancer: the Atherosclerosis Risk in Communities (ARIC) Study. *Diabetes Care.* 2008; 31: 905-9.
- 19) Samani NJ, Erdmann J, Hall AS, et al. Genome-

- wide association analysis of coronary artery disease. *N Engl J Med.* 2007; 357: 443-53.
- 20) Cookson W, Liang L, Abecasis G, et al. Mapping complex disease traits with global gene expression. *Nat Rev Genet.* 2009; 10: 184-94.
 - 21) Kottgen A, Glazer NL, Dehghan A, et al. Multiple loci associated with indices of renal function and chronic kidney disease. *Nat Genet.* 2009; 41: 712-7.
 - 22) Kottgen A, Pattaro C, Boger CA, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 2010; 42: 376-84.
 - 23) Chambers JC, Zhang W, Lord GM, et al. Genetic loci influencing kidney function and chronic kidney disease. *Nat Genet.* 2010; 42: 373-5.
 - 24) Pezzolesi MG, Poznik GD, Mychaleckyj JC, et al. Genome-wide association scan for diabetic nephropathy susceptibility genes in type 1 diabetes. *Diabetes.* 2009; 58: 1403-10.
 - 25) Shimazaki A, Kawamura Y, Kanazawa A, et al. Genetic variations in the gene encoding ELMO1 are associated with susceptibility to diabetic nephropathy. *Diabetes.* 2005; 54: 1171-8.
 - 26) Maeda S, Kobayashi MA, Araki S, et al. A single nucleotide polymorphism within the acetyl-coenzyme A carboxylase beta gene is associated with proteinuria in patients with type 2 diabetes. *PLoS Genet.* 2010; 6: e1000842.
 - 27) Feehally J, Farrall M, Boland A, et al. HLA has strongest association with IgA nephropathy in genome-wide analysis. *J Am Soc Nephrol.* 2010; 21: 1791-7.
 - 28) Akiyama F, Tanaka T, Yamada R, et al. Single-nucleotide polymorphisms in the class II region of the major histocompatibility complex in Japanese patients with immunoglobulin A nephropathy. *J Hum Genet.* 2002; 47: 532-8.
 - 29) Chi KR. Human genetics: Hit or miss? *Nature.* 2009; 461: 712-4.
 - 30) Visscher PM. Sizing up human height variation. *Nat Genet.* 2008; 40: 489-90.
 - 31) Weedon MN, Lettre G, Freathy RM, et al. A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat Genet.* 2007; 39: 1245-50.
 - 32) Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009; 461: 747-53.
 - 33) Cohen JC, Kiss RS, Pertsemlidis A, et al. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science.* 2004; 305: 869-72.
 - 34) Romeo S, Pennacchio LA, Fu Y, et al. Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. *Nat Genet.* 2007; 39: 513-6.
 - 35) Weiss LA, Shen Y, Korn JM, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med.* 2008; 358: 667-75.
 - 36) International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature.* 2008; 455: 237-41.
 - 37) Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. *Nature.* 2008; 455: 232-6.
 - 38) Bassett AS, Scherer SW, Brzustowicz LM. Copy number variations in schizophrenia: critical review and new perspectives on concepts of genetics and disease. *Am J Psychiatry.* 2010; 167: 899-914.
 - 39) McCarroll SA, Huett A, Kuballa P, et al. Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet.* 2008; 40: 1107-12.
 - 40) de Cid R, Riveira-Munoz E, Zeeuwen PL, et al. Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat Genet.* 2009; 41: 211-5.
 - 41) Moore JH, Asselbergs FW, Williams SM. Bioinformatics challenges for genome-wide association studies. *Bioinformatics.* 2010; 26: 445-55.
 - 42) Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet.* 2010; 11: 446-50.
 - 43) Moore JH, Williams SM. Epistasis and its implications for personal genetics. *Am J Hum Genet.* 2009; 85: 309-20.
 - 44) Tishkoff SA, Reed FA, Friedlaender FR, et al. The genetic structure and history of Africans and African Americans. *Science.* 2009; 324: 1035-44.
 - 45) Teo YY, Small KS, Kwiatkowski DP. Methodological challenges of genome-wide association analysis in Africa. *Nat Rev Genet.* 2010; 11: 149-60.

Annual Review ^{じんぞう}腎臓 2011 ©

発行 2011年 1月25日 初版 1刷

編集者 富野 康日己
柏原 直樹
成田 一衛

発行者 株式会社 中外医学社
代表取締役 青木 滋

〒162-0805 東京都新宿区矢来町 62
電 話 03-3268-2701 (代)
振替口座 00190-1-98814 番

印刷 / 東京リスマチック (株) < HI・YT >
製本 / 田中製本 (株) Printed in Japan
ISBN978-4-498-12472-1

JCOPY < (社) 出版者著作権管理機構 委託出版物 >

本書の無断複写は著作権法上での例外を除き禁じられています。
複写される場合は、そのつど事前に、(社)出版者著作権管理機構
(電話 03-3513-6969, FAX 03-3513-6979, e-mail: info@jcopy.
or.jp) の許諾を得てください。