

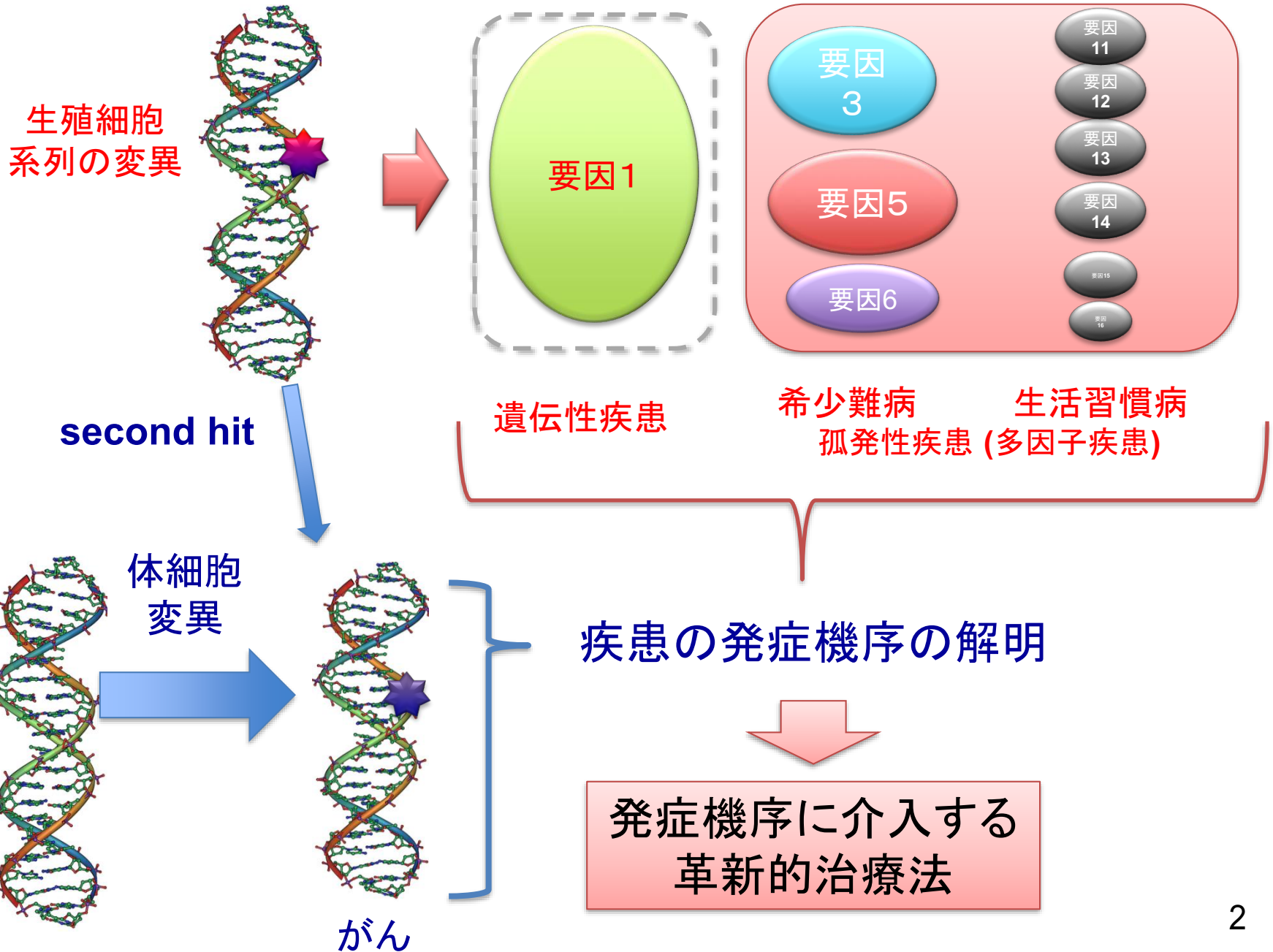
孤発性疾患（難病）の発症原因の解明・ 治療法実現に向けて何をなすべきか？

東京大学大学院医学系研究科 分子神経学

国際医療福祉大学ゲノム医学研究所

辻 省次

ゲノムサイエンスから見た疾患の発症機構



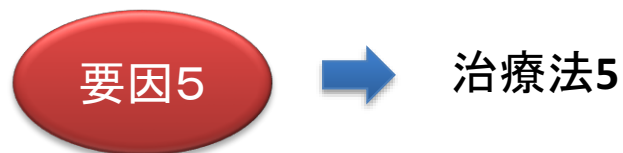
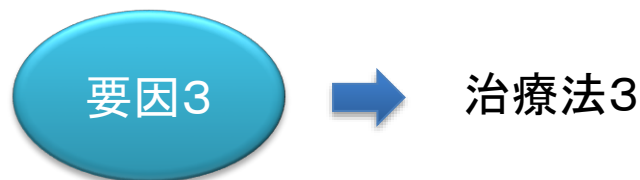
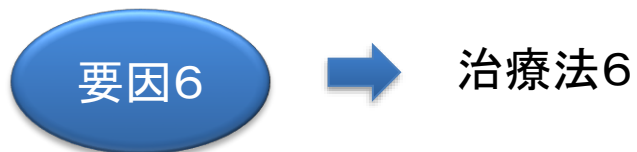
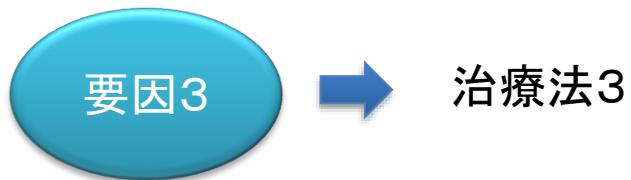
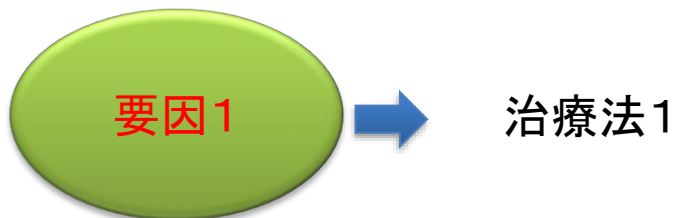
孤発性疾患（多因子疾患）の原因となるゲノム要因と治療法



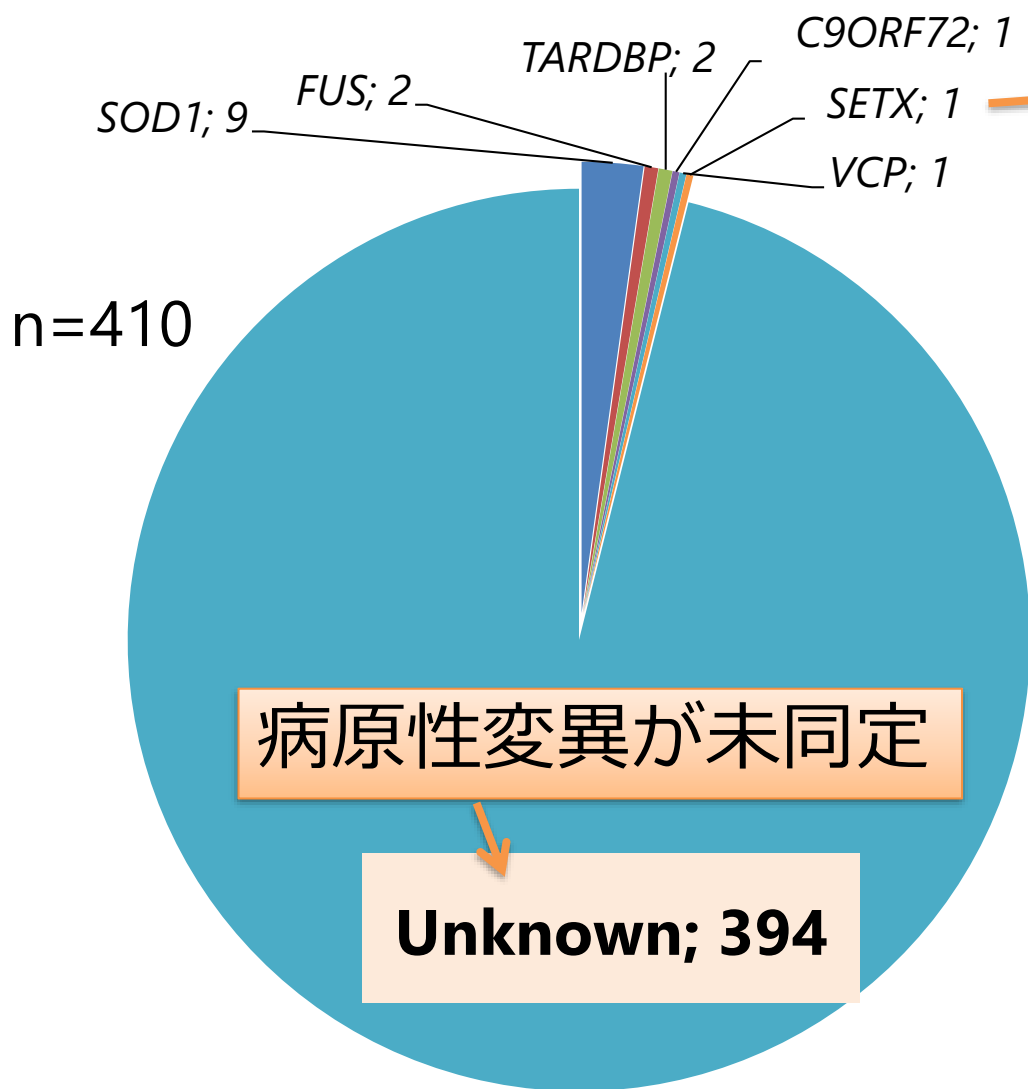
Aさん



Bさん

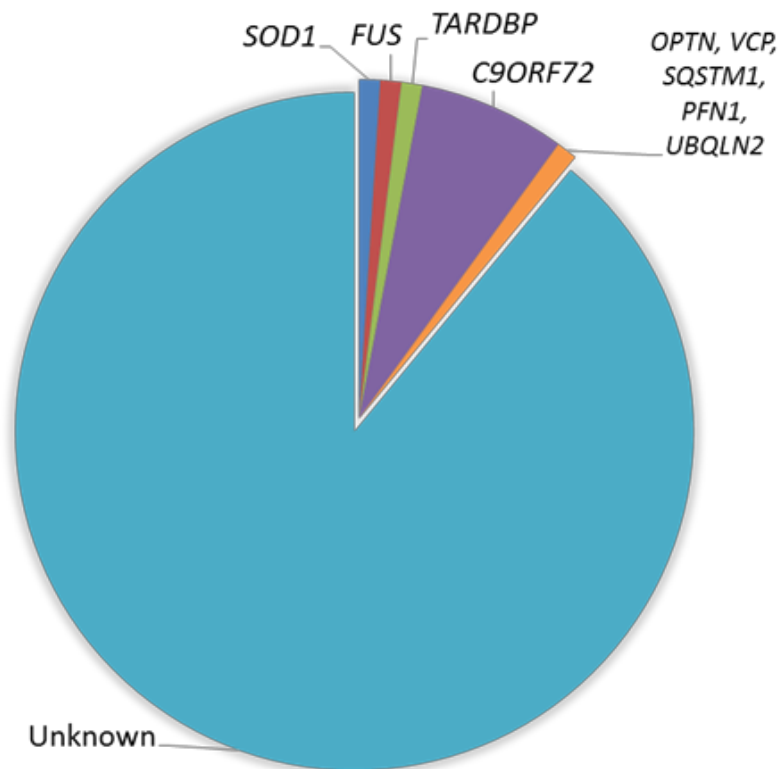


孤発性ALSの分子疫学



当科 (日本人集団)

SALSの16例 (4%)
にALSの病因遺伝子
の病原性変異を検出



欧州コホート*

4

パーキンソン病のリスク遺伝子：アレル頻度とオッズ比

Variants (SNPs)	パーキンソン病	コントロール	オッズ比 (95% 信頼区間)
<i>GBA</i> *	9.4%	0.4%	28.0 (7.3 to 238.3)
<i>SNCA</i> (rs11931074)**	36%	42%	1.37 (1.27–1.48)
<i>LRRK2</i> (rs1994090)**	10%	8%	1.39 (1.24–1.56)
<i>BST1</i> (rs11931532)	47 %	42%	1.24 (1.15–1.33)
<i>PARK16</i> (rs947211)**	42%	50%	1.30 (1.21–1.39)

- 影響度が大きいと、治療標的になる可能性が高い
- *GBA*については、基質合成阻害薬を用いた、パーキンソン病に対する企業治験が進行中。

* Mitsui et al. Arch. Neurol (2010)

** Satake et al. Nature Genet (2009)

低頻度アレルの中に、高いodds ratioを示す遺伝子が見出されてきている

リスク

100

10

1

0.01

0.1

1.0

10

100 %

企業主導治験

Gaucher病の治療薬

Rare variants

基質合成阻害剤

GBA

(パーキンソン病)

Rare variants in COQ2

(多系統萎縮症)

医師主導治験

高用量CoQ10の補充

V393A in COQ2

(多系統萎縮症)

ε4 allele
in APOE
in AD

基質合成経路の阻害薬

グルコセレブロシド

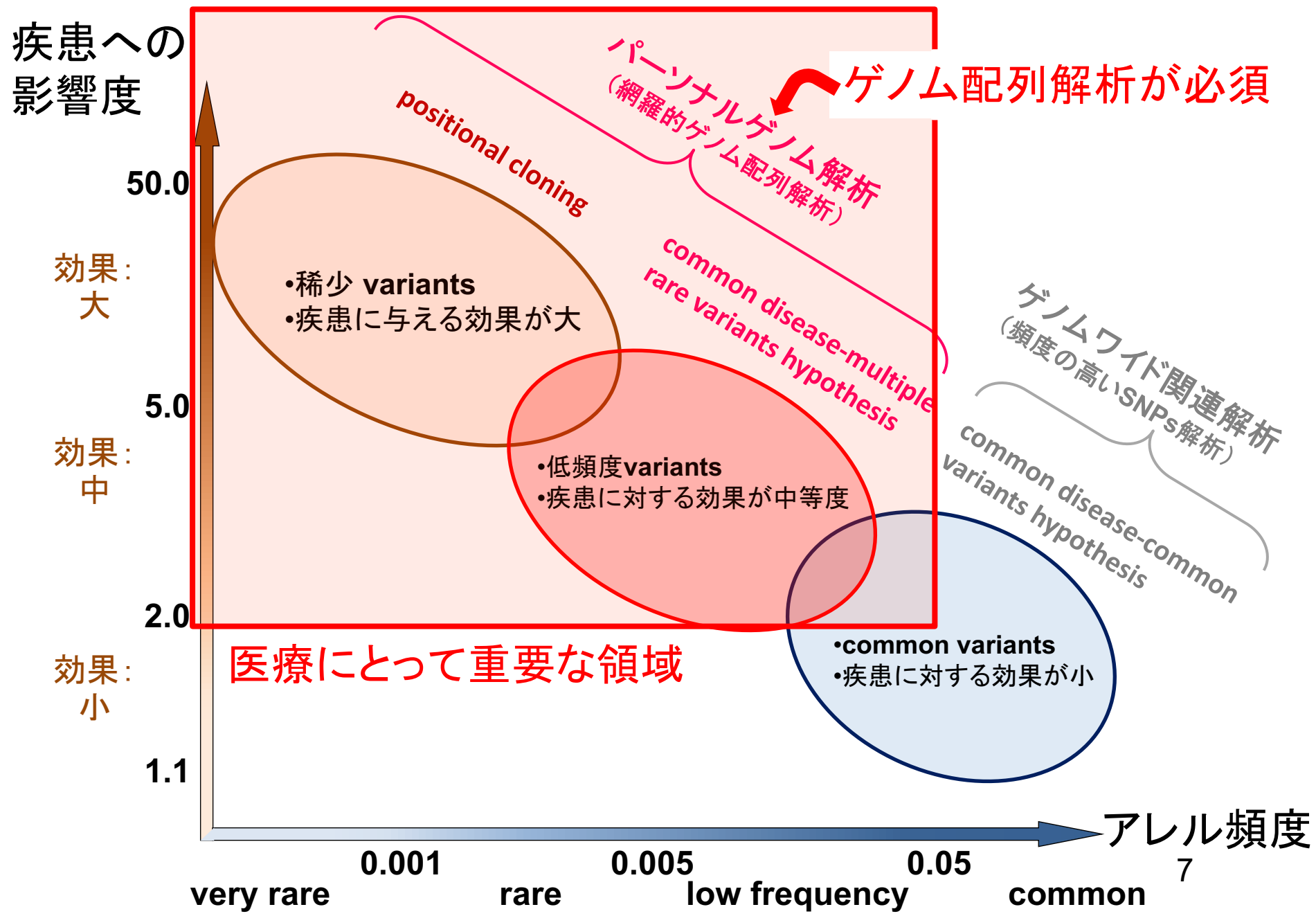
分解する酵素:
グルコセレブロシダーゼ
遺伝子: GBA

セラミド

GWASによって見出された疾患
感受性遺伝変異の疾患発症に
対する影響度は小さい

集団におけるアレル頻度

Missing heritabilityを解くことが孤発性疾患の発症機序解明の鍵



Recovery of trait heritability from whole genome sequence data.

P. Wainschtein et al. Preprint at bioRxiv
<https://doi.org/10.1101/588020>; 2019

GENETICS

Height's 'missing heritability' found

Analysis of 20,000 genomes reveals that rare gene variants can help to explain how the trait is inherited.

Linda Geddes. *Nature* 568: 444, 2019

背景

臨床遺伝学で求められていたヒトの伸長に関するheritability: **~0.8**

GWASの成果と限界

GWASによって求められたheritability (h^2_{SNP}) = **0.45** (0.08)

Yang et al. *Nature Genetics* 42: 565, 2010



missing heritability

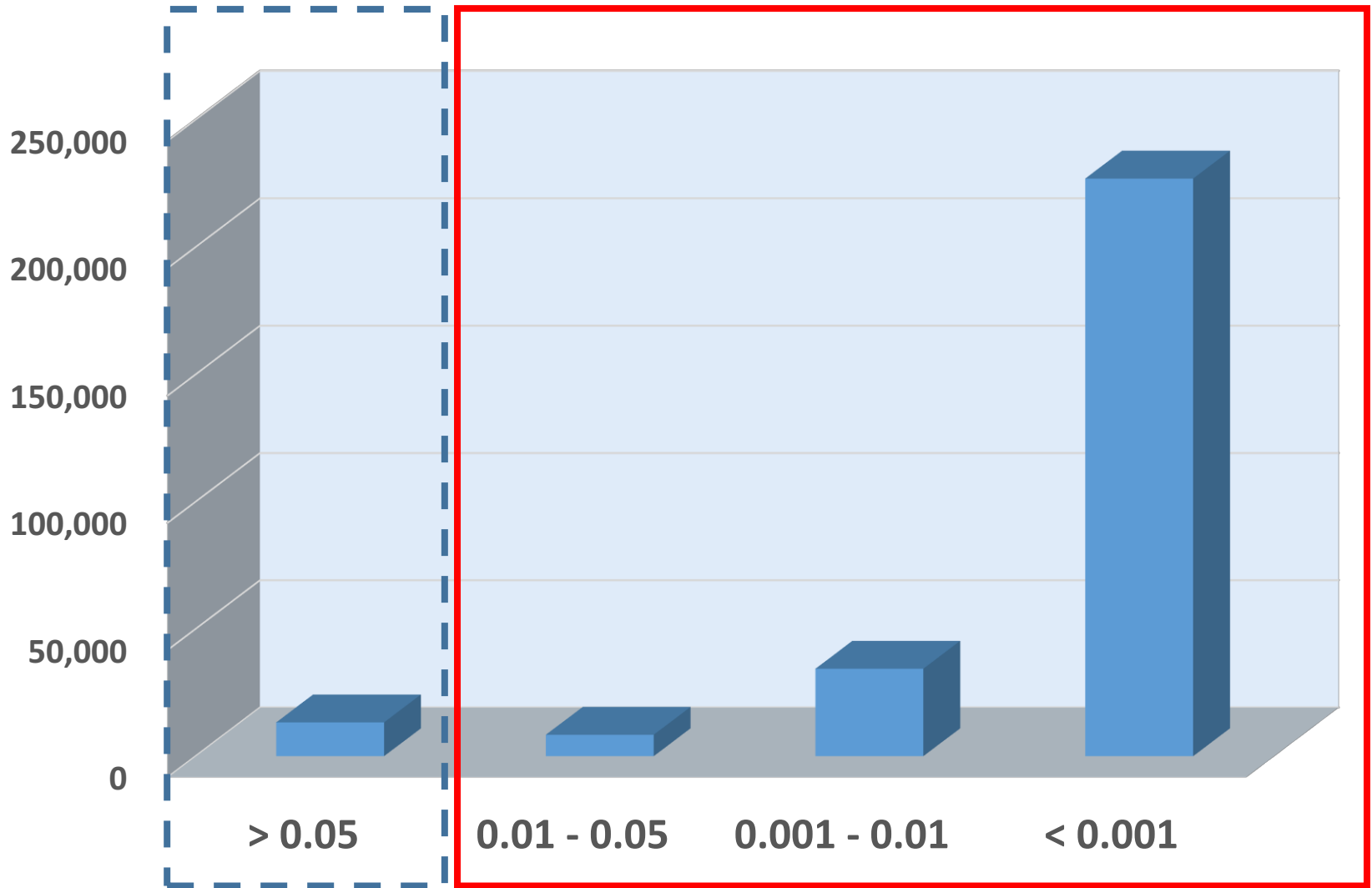
21,620人の全ゲノムシーケンスデータに基づき, 身長とBMIのheritabilityを求めた.

- heritability of height: **0.79** (SE 0.09)
- MAF (minor allele frequency) が, 0.0001-0.1 の範囲のrare variantsが, heritability として, **0.54** (SE 0.05) を説明できる.

ゲノムシーケンス解析によって見出されるvariantsの大部分はrare variantsである

GWAS

本当に解析が必要な領域



- トリオ解析(両親とその子供)を解析すると、新生突然変異(非同義置換)が1-2個/世代で生じている。
→日本では100-200万個の新生非同義置換が毎年生じている。
- 農耕文明の定着に伴い生じた人口爆発により、各地域に特異的な低頻度変異が集積している。

Searching for missing heritability: Designing rare variant association studies

Or Zuk^{a,b,1}, Stephen F. Schaffner^a, Kaitlin Samocha^{a,c,d}, Ron Do^{a,e}, Eliana Hechter^a, Sekar Kathiresan^{a,e,f,g}, Mark J. Daly^{a,c}, Benjamin M. Neale^{a,c}, Shamil R. Sunyaev^{a,h}, and Eric S. Lander^{a,i,j,2}

PNAS E455-E464 (2014)

GWAS = common variant association study (CVAS)



missing heritabilityを解決するためのパラダイムシフト

rare variant association study (RVAS)

RVASの研究パラダイム

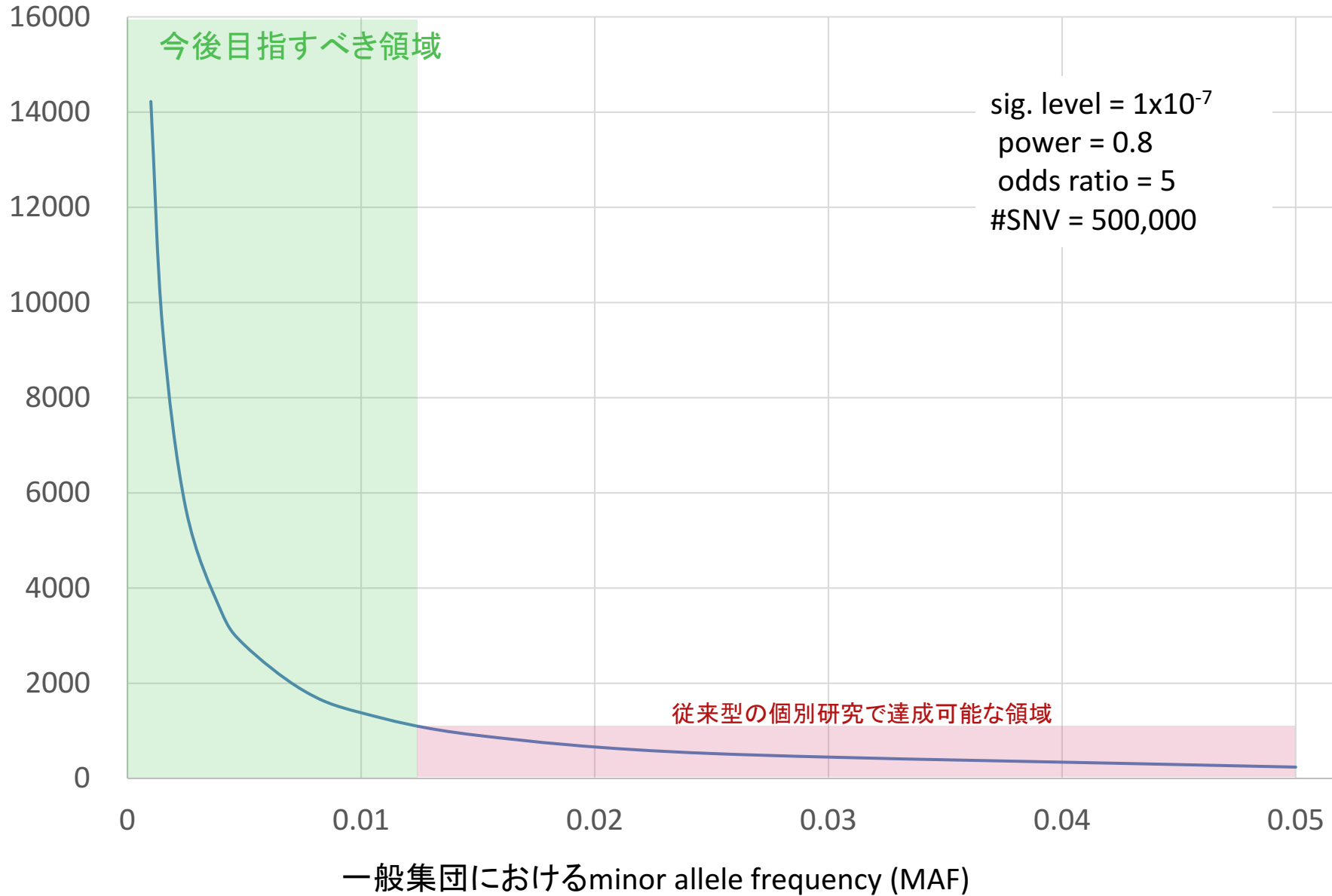
whole exome sequence/whole genome sequence解析に基づき、ゲノム上のすべてのvariantsを検出して、患者集団において、有意に高頻度に観察されるvariantsを検出し、そのvariantsが存在する遺伝子を同定、疾患発症の分子基盤を明らかにする。



- 見出されるゲノム上のvariantsが膨大な数になるために、多重検定を考慮に入れた、ゲノムワイドの検出力を確保するためには、**十分に大きなサンプルサイズ**が必要となる。
- 十分な検出力を確保した研究デザインで、RVASを実施する必要がある。これが満たされないと、type 2 error (第2種過誤)のために、under-powerの結果となる。

検出力=0.8を確保するために必要なサンプル数

rare variantsになるほど、解析に必要なサンプル数が増大する



難病の原因究明を目指した全ゲノムシーケンス解析の研究パラダイム

背景

全ゲノムシーケンス解析から得られるゲノム上のvariantsの数は膨大であり(1人300万個程度), これらのvariantsの中で, 疾患発症に関与するvariantsは少数で, 大部分は中立のvariantsであると考えられる(一部にはprotectiveに作用するものもあるかもしれない).



考慮すべき事

- 検出されたvariantsについて, 患者群と健常者群で, その頻度分析をする場合に, 多重検定を考慮に入れて十分な検出力を達成するためには, 十分なサンプルサイズが必須となる. これが実現できていないと, type 2 error (第2種過誤)に陥る(under-power).
- 従って, いかに適切なサンプルサイズを設計するかが最重要課題となる.

戦略1: 対象疾患に関して

1. 対象疾患の検討と, 必要とされるサンプルサイズの検討
2. コホート研究に基づく, 臨床情報, バイオリソースの集積(ゲノムデータとの統合)
3. バイオリソース, 臨床情報, ゲノム情報の幅広い利活用を可能にするインフォームドコンセントの取得. Securityの高い管理体制
4. 比較対象となる, コントロール群の集積
5. replication setの準備(国際共同研究も視野に含める)

戦略2: RVASの研究パラダイム

1. 必要とされるサンプルサイズの検討.
2. 大規模ゲノム解析拠点
3. variantsの絞り込み(allele frequency, 非同義置換, 機能予測, 生物学的なpathway)
4. 遺伝統計学に基づく解析アルゴリズムの研究を含めた解析拠点
5. 人材育成を含めた, オープンなゲノム医学研究拠点

戦略1:対象疾患に関して

- 指定難病の中で、患者数が圧倒的に多いのは、孤発性疾患であり、全ゲノムシーケンス解析が、その病因解明に大きく寄与すると考えられる。(例:パーキンソン病の患者数を15万人とすると、遺伝性パーキンソン病は5-10%とされており、孤発性パーキンソン病は13.5万人と推定される)
- 神経変性疾患と位置づけられるものは、臨床的に同一の疾患と位置づけられるが、その中に、遺伝性、孤発性の疾患が含まれることが特徴。臨床所見だけでなく、病理学的所見からも共通するところが多く、孤発性疾患においても遺伝的な基盤が想定され、遺伝統計学的に基づく研究からも、一定のheritabilityが認められる疾患が多い。
- 全ゲノムシーケンスの対象疾患としては、指定難病の中で、heritability (遺伝率)が高いことが示されている疾患を対象とするのが良いのではないか。
- 上記の考察からは、パーキンソン病、筋萎縮性側索硬化症、脊髄小脳変性症、多系統萎縮症などが考慮される。
- 指定難病の中で、自己免疫疾患と分類される疾患は、別途、免疫学、免疫遺伝学の面から、その研究パラダイムについての検討が必要であるが、炎症性腸疾患 (Crohn病、潰瘍性大腸炎では、高いheritabilityが報告されており、そのheritabilityを十分に説明できるだけの遺伝子が見いだされていないとされており、解析対象として検討する)。

難病のheritability

パーキンソン病

- 8 cohorts (7,096 case and 19 455 controls of European ancestry)
- Genome-wide complex trait analysis (GCTA): 全てのSNPsを投入してheritabilityを推定.

	Heritability estimate (meta-analysis, random effects), (95%CI)
All SNPs	0.24 (0.16 - 0.31)
GWAS SNPs in PD loci	0.02 (0.01 - 0.03)
Non-GWAS region	0.23 (0.14 - 0.32)

Keller et al. *Hum Mol Genet* 21: 4996-5009 (2012)

ALS (筋萎縮性側索硬化症)

ALSの発症が観察された双生児(171組)を研究対象. 孤発性ALSの遺伝率 (heritability) は, **0.61** (0.38-0.78)と推定された.

Al-Chalabi et al. *J Neurol Neurosurg Psychiatry* 81:1324-1326 (2010).

炎症性腸疾患 (Crohn's disease, Ulcerative colitis)

pooled twin studies に基づく, heritabilityの推定

Crohn's disease : **0.75**, Ulcerative colitis = **0.67**

Gordon et al *Inflamm Bowel Dis* 21: 1428-34, (2015)

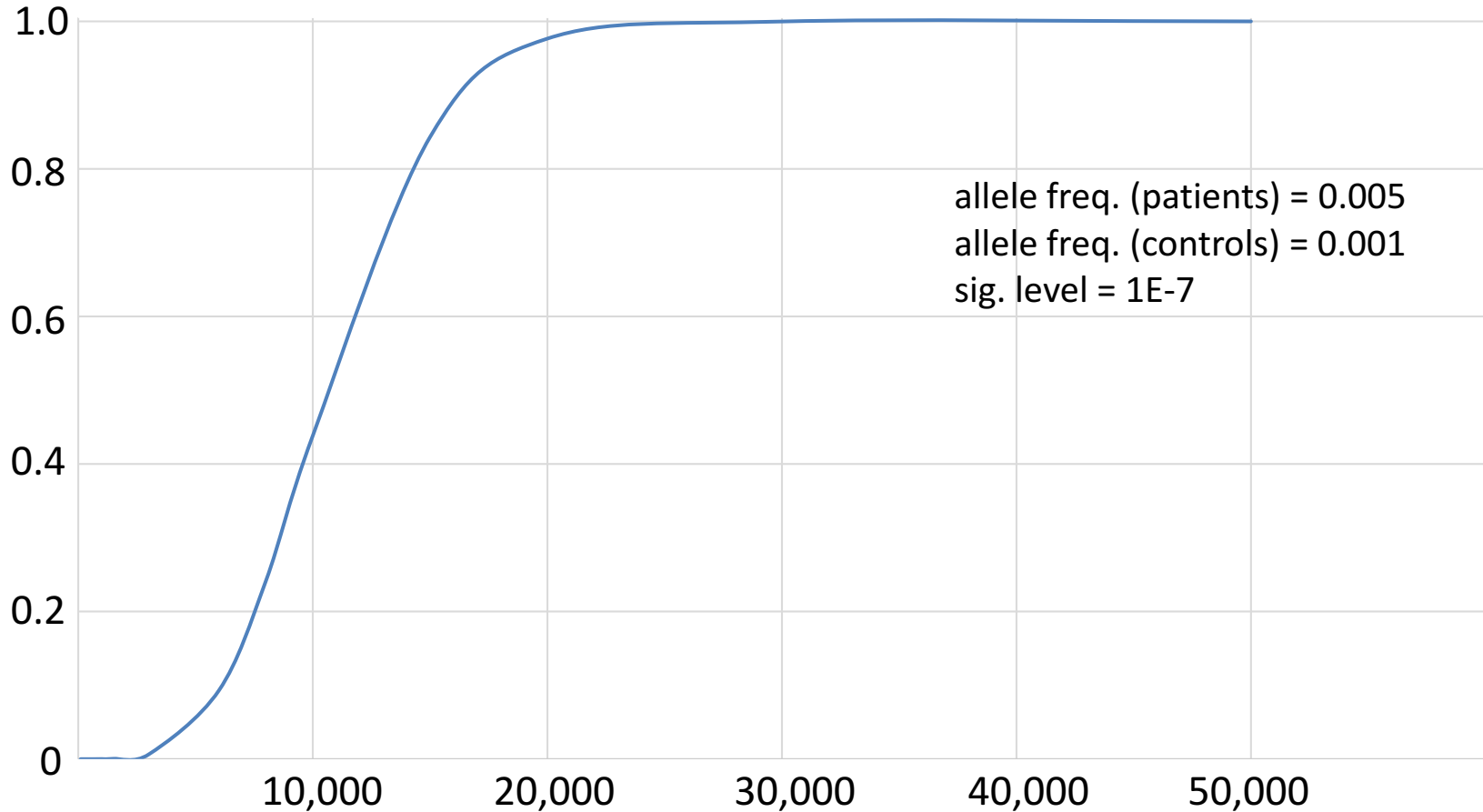
わが国における、難病克服を目指した全ゲノムシーケンスプロジェクト

1. 患者数が圧倒的に多い難病の発症原因の解明と治療法開発を目指した研究を推進する。
2. 難病の発症原因の解明は、heritabilityが高いことが示されている疾患で、患者数の多い疾患に集中する。
3. 難病の発症機構(例えば、神経変性疾患、自己免疫疾患など)に応じて、適切なプロジェクトを設定する。
4. 大規模全ゲノムシーケンスプロジェクトの推進には、次の要素が必須である。
 - コホート研究に基づく、症例の蓄積、臨床情報、バイオリソースの集積、データの共有(オールジャパン体制が基本)
 - 解析規模としては、1疾患当たり、25,000程度を目指す(odds ratioとして、3以上、アレル頻度として、0.02-0.001程度)。目標到達までに、段階を踏む戦略立案が現実的。
例:①現状で、1,000例以上の集積ができていないコホートについて、直ちに全ゲノムシーケンスを実施。
②並行して、新たなコホートの立ち上げ。①、②で1疾患当たり、5,000例の集積を目指す。
③1疾患あたり、5,000例の全ゲノム解析を目指し、到達時点評価、次のステップの戦術を検討。
 - 大規模全ゲノムシーケンス拠点の形成。
 - ゲノムインフォマティクス、遺伝統計学研究拠点の形成。(単に全ゲノムシーケンスをするだけでは、不十分。遺伝統計学を含めた研究体制の整備、人材育成、臨床系の研究者の研究参加体制の整備が必須)
 - 拠点は、できるだけオープンな設計にして、研究者が幅広く参加できること、データの積極的な利活用ができること
 - 情報、バイオリソース、ゲノムデータの共有、利活用を可能にするインフォームドコンセントの取得が前提条件
 - Public-private partnership (PPP) に基づく、企業の積極的な参加

參考資料

サンプルサイズを大きくすれば, 検出力は, 1 に近づく

Power



サンプルサイズ(人)

全ゲノムシーケンス vs. 全エクソンシーケンス

全ゲノムシーケンスがなぜ必要か？

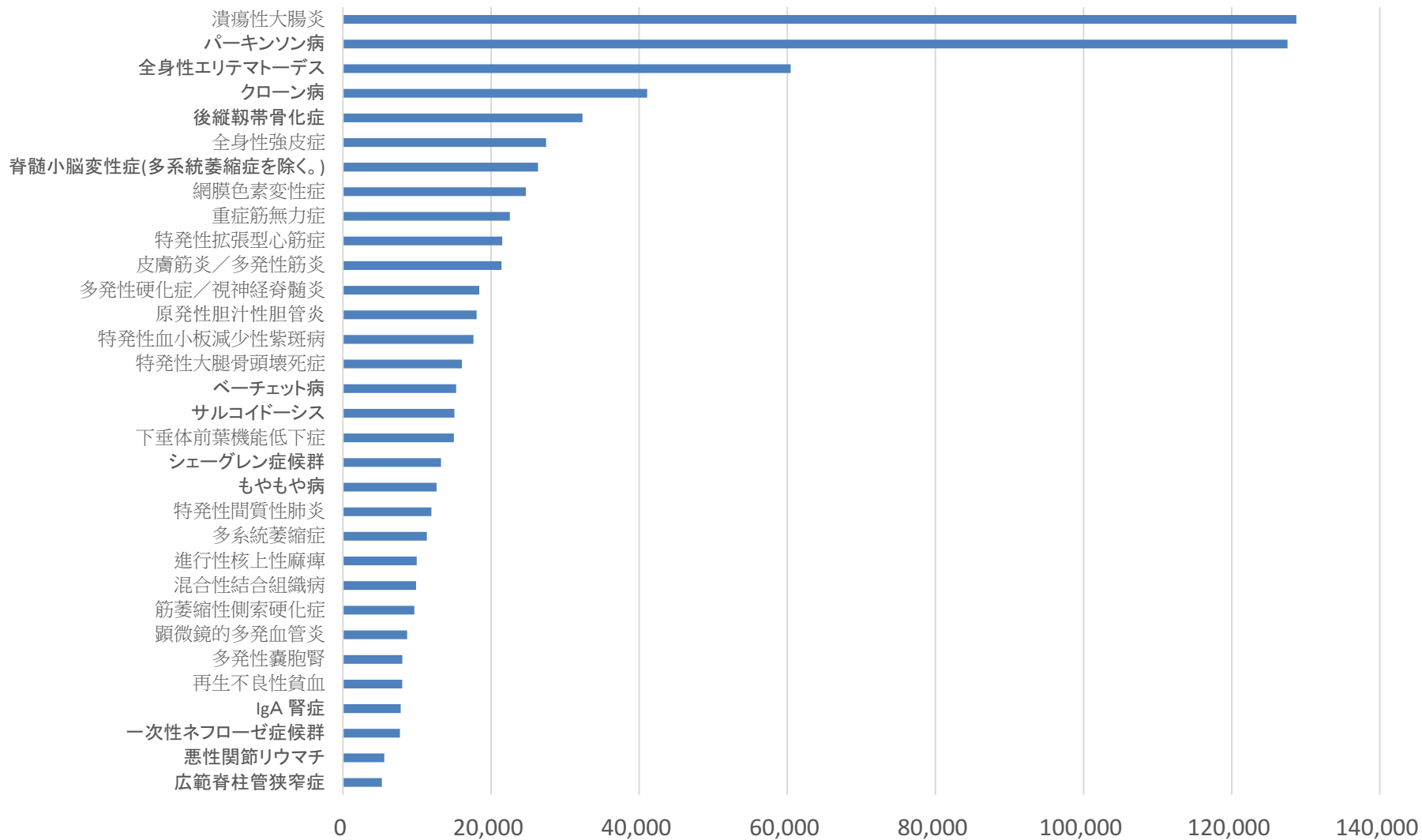
Pros

- 翻訳領域においても、全エクソンシーケンスに比較して、均一なcoverageが得られる。
- 全エクソンシーケンスは、エクソン領域の濃縮のためにプローブデザインが必要（annotationがされていない遺伝子は、captureできない。例：神経核内封入体病の原因遺伝子NBPF19は、hg38で初めてannotationされた）
- 翻訳領域以外の領域を含め、ゲノム上のすべてのvariantsが検出できる。
- 構造変異，コピー数変異，反復配列の異常伸長などは，全エクソンシーケンスでは検出が困難。
- 全ゲノムシーケンスデータは，将来的に，様々な利活用できるという点で，疾患のゲノム研究の発展に大きく寄与する。

Cons

- 全ゲノムシーケンスの費用は，低減してきているとはいえ，全エクソンシーケンスよりも高額である。将来的には，さらにコストの低減が期待される。
- ゲノムインフォマティクス解析の負荷，data storageの負担が大きい。
- エクソン以外の領域以外で，プロモーター領域，エンハンサー領域などは疾患発症に一定の関与があると考えられるが，それ以外の広大なゲノム領域が，どの程度疾患発症に関与しているのか，明らかでない。

指定難病においては、孤発性疾患が大部分を占める



疾患関連遺伝子の探索方法（ゲノムワイド関連解析, GWAS）

創始者の染色体



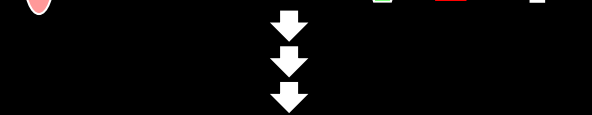
疾患感受性遺伝子

多型性

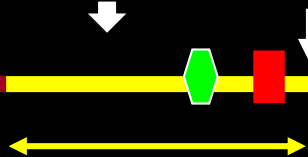
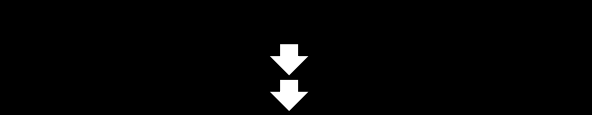
組換え



疾患関連遺伝子と連鎖不平衡にある多型locusは、疾患関連遺伝子に近いほど、集団において、疾患関連遺伝子に付随して観察される



病気の人の集団は  を高頻度に有している



この連鎖不平衡領域を代表するcommon SNP (MAF > 0.05)  を用いて、患者集団と健常者集団で頻度の比較をする

戦略2: RVASの研究パラダイム

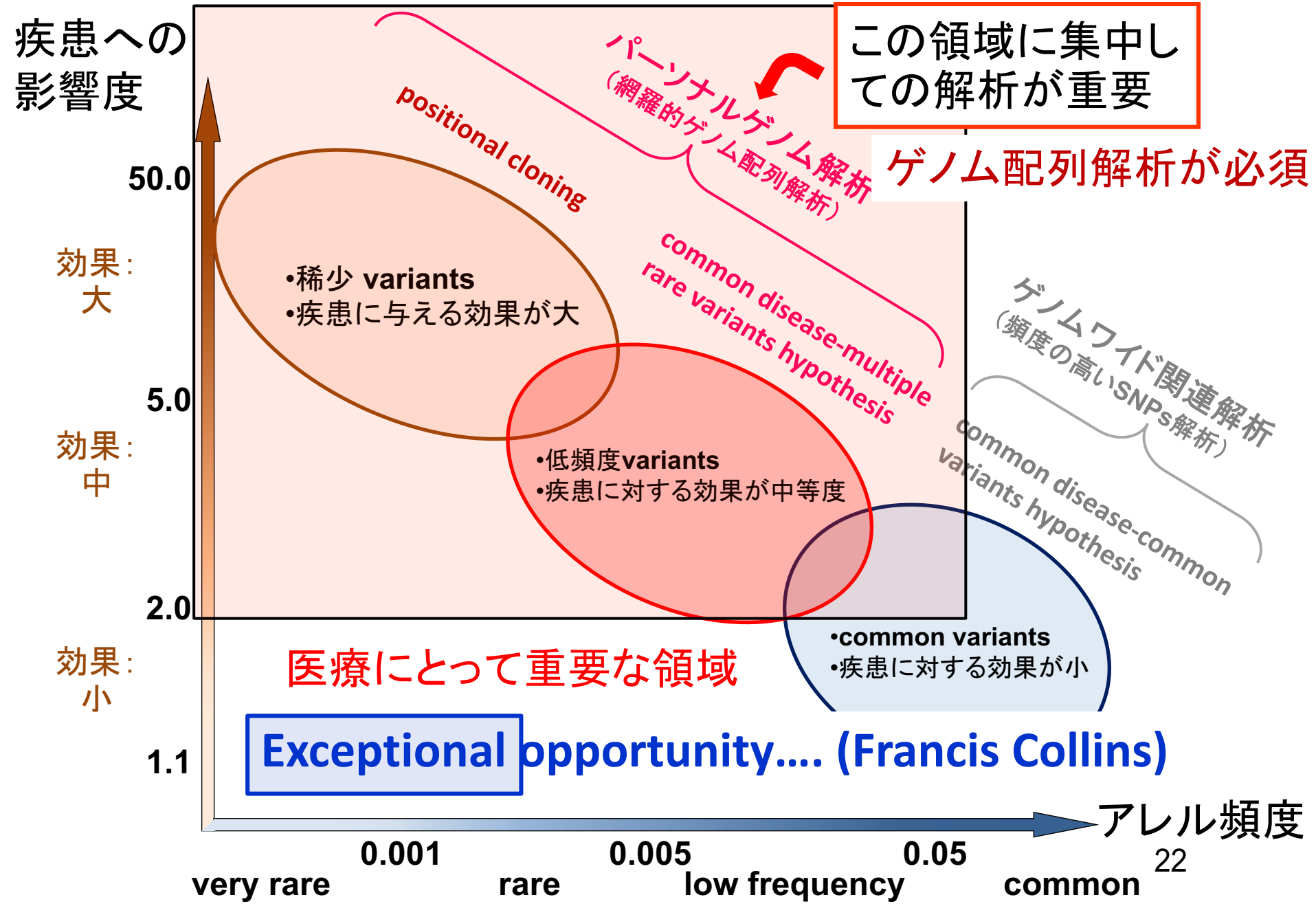
1. 必要とされるサンプルサイズの評価.
2. variantsの絞り込み (allele frequency, 非同義置換, 機能予測, 生物学的なpathway)
3. 遺伝統計学に基づく解析アルゴリズム

例:

Larson et al. gsSKAT: Rapid gene-set analysis and multiple testing correction for rare-variant association studies using weighted linear kernels. *Genet Epidemiol* (2017)

- 検出されるvariantsを全て統計解析に投入すると, 巨大なサンプルサイズが必要になってしまう.
- variantsを一定のルールに基づき, フィルターをすることが現実的 (非同義置換, rare variantsであること, loss of functionのvariants, 機能予測からdeleterious variantsと推定される. . .)

精神・神経疾患遺伝子探索の新しいパラダイム



アレル頻度とEffect sizeに基づく疾患関連遺伝子探索のパラダイム

