

希少・未診断疾患イニシアチブにおける 大規模全エクソーム解析の 俯瞰的分析

国立成育医療研究センター
ゲノム医療研究部
佐藤万仁、柳久美子、要匡

所属機関紹介

名称：国立成育医療研究センター

所在地：東京都世田谷区

沿革：厚生労働省所管の
6つの国立高度専門医療研究センター（NC）の1つ

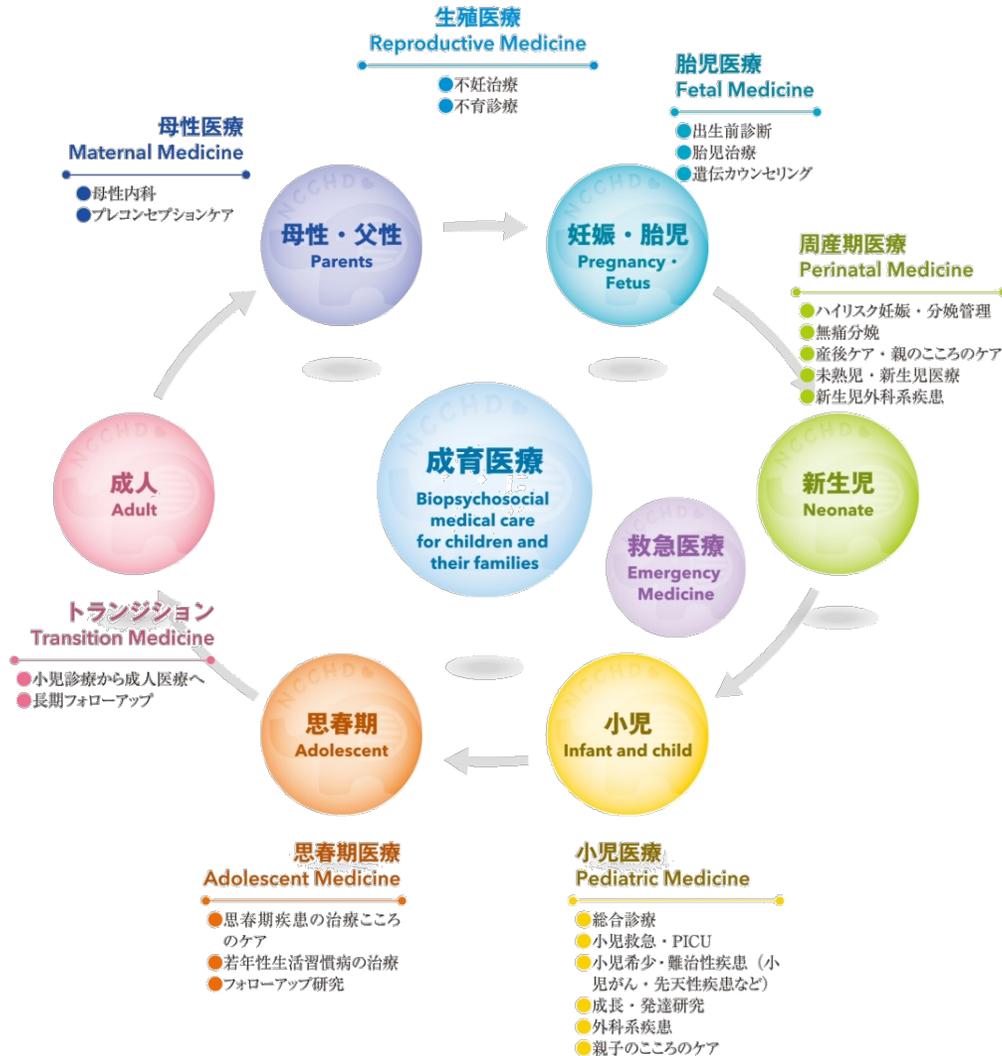
組織：病院と研究所を併設

理念：成育医療の提供と成育疾患の研究



成育医療・成育疾患

成育医療



成育疾患

- 成育疾患には遺伝的要因が大きく関与
- 希少・難治性疾患は6,000以上知られているが、うち80%は遺伝性疾患
- 遺伝性疾患の50%は病因遺伝子が不明
- 本邦では2019年7月1日現在、333の疾患が指定難病として告示

ゲノム医療

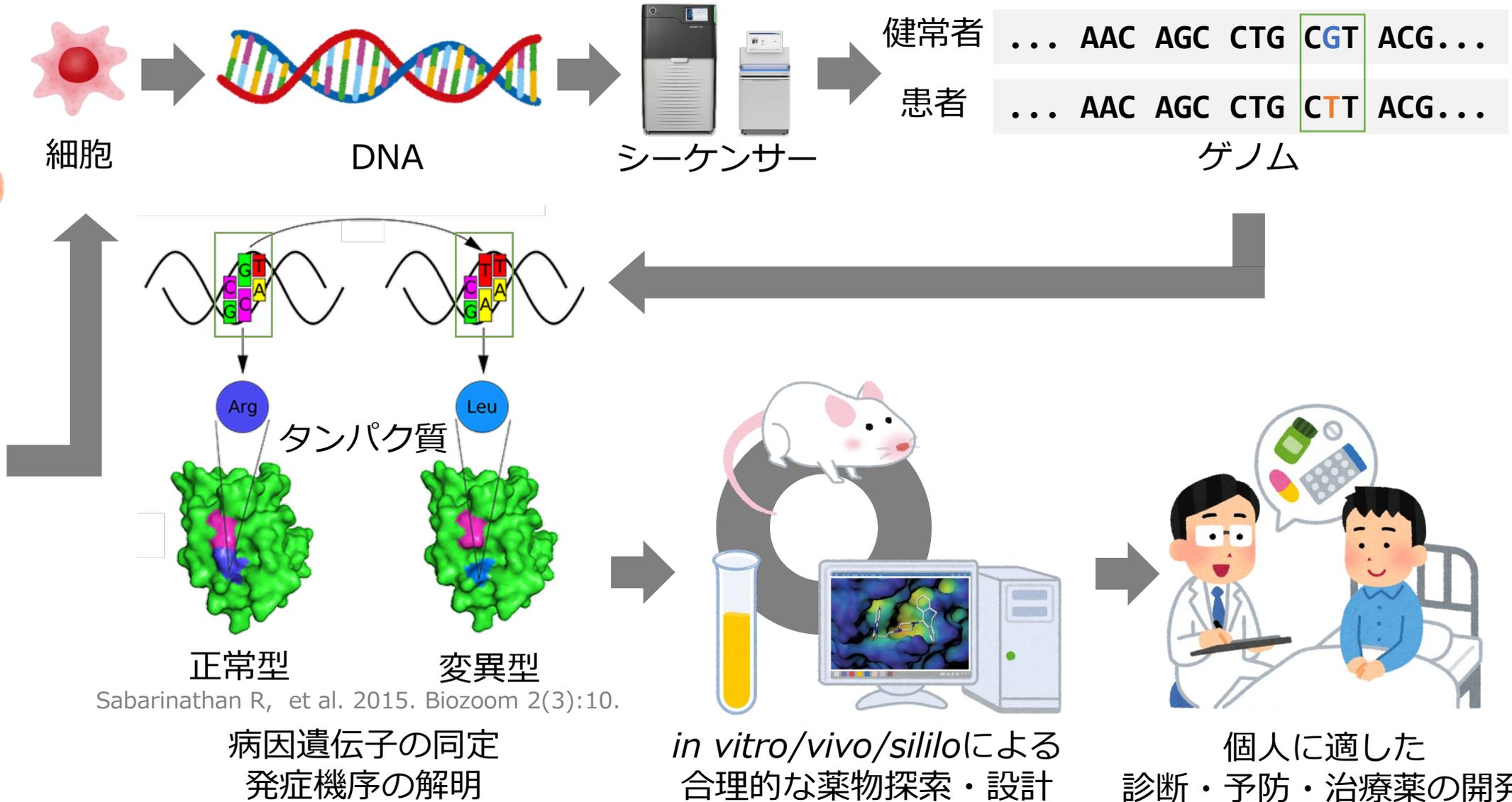
- 患者や家系の遺伝情報（ゲノム）を解読・比較
- 遺伝性疾患の病因や機序を解明
- 診断・治療・予防の手段や方法を開発

ゲノム医療



ヒト

2020-05-21



Sabarinathan R, et al. 2015. Biozoom 2(3):10.

難治性疾患：指定難病の例

疾患数	関連する症候群	受給者証保持者数（2017年度）
333		892,445
83	神経・筋疾患	259,052
20	消化器系疾患	194,967
26	免疫系疾患	146,386
13	骨・関節系疾患	61,492
15	皮膚・結合組織疾患	39,930
14	呼吸器系疾患	34,883
13	血液系疾患	29,452
21	循環器系疾患	26,990
16	内分泌系疾患	26,179
13	腎・泌尿器系疾患	25,586
8	視覚系疾患	24,899
1	皮膚・結合組織疾患、免疫系疾患	9,871
43	代謝系疾患	6,721
1	免疫系疾患、耳鼻科系疾患	4,978
42	染色体または遺伝子に変化を伴う症候群	1,011
2	耳鼻科系疾患	32
1	視覚系疾患、耳鼻科系疾患	10
1	聴覚・平衡機能系疾患	6

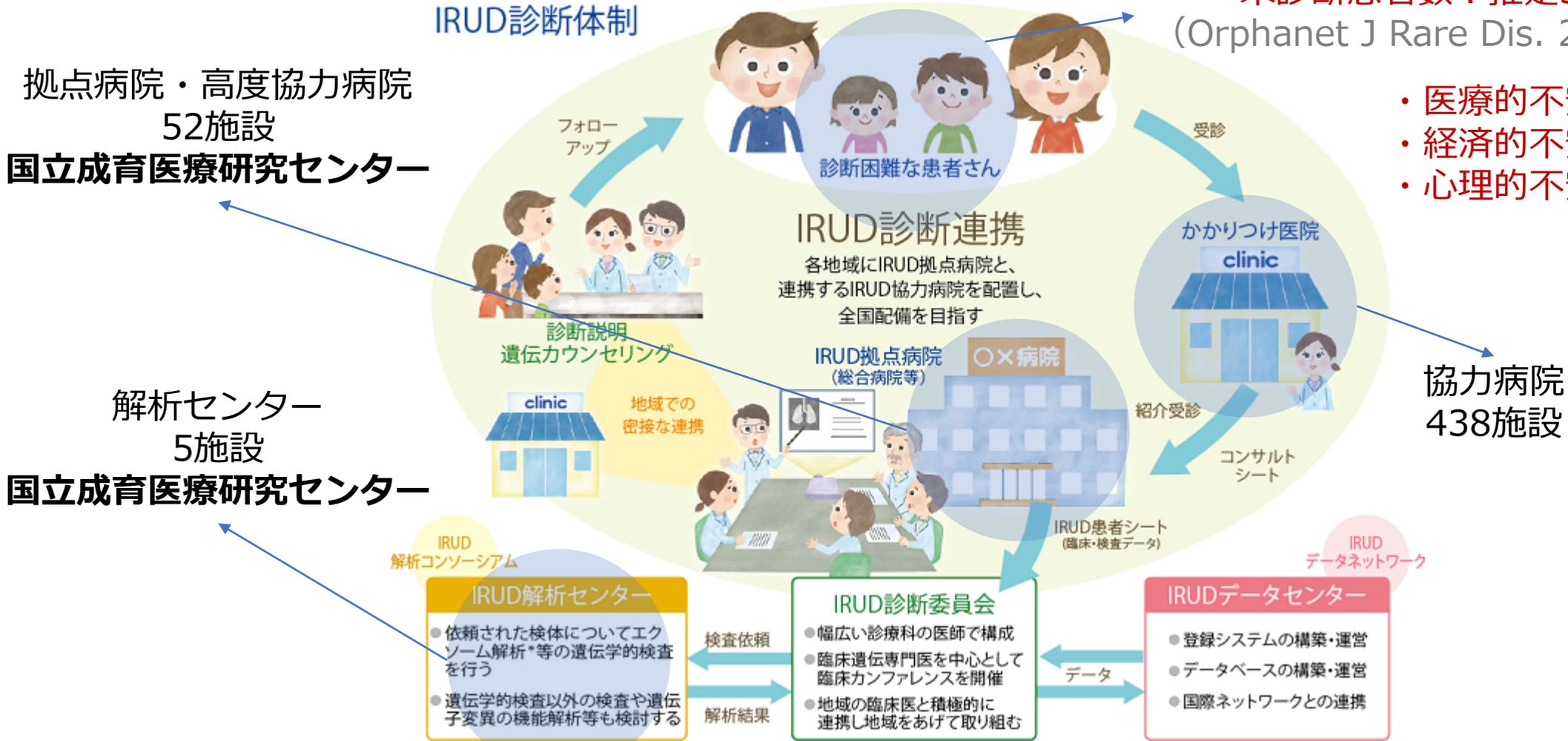
- ・筋萎縮性側索硬化症（ALS）
- ・パーキンソン病
- ・もやもや病
- ・筋ジストロフィー

本邦総人口の
0.70%

未診断疾患イニシアチブ Initiative on Rare and Undiagnosed Diseases (IRUD)

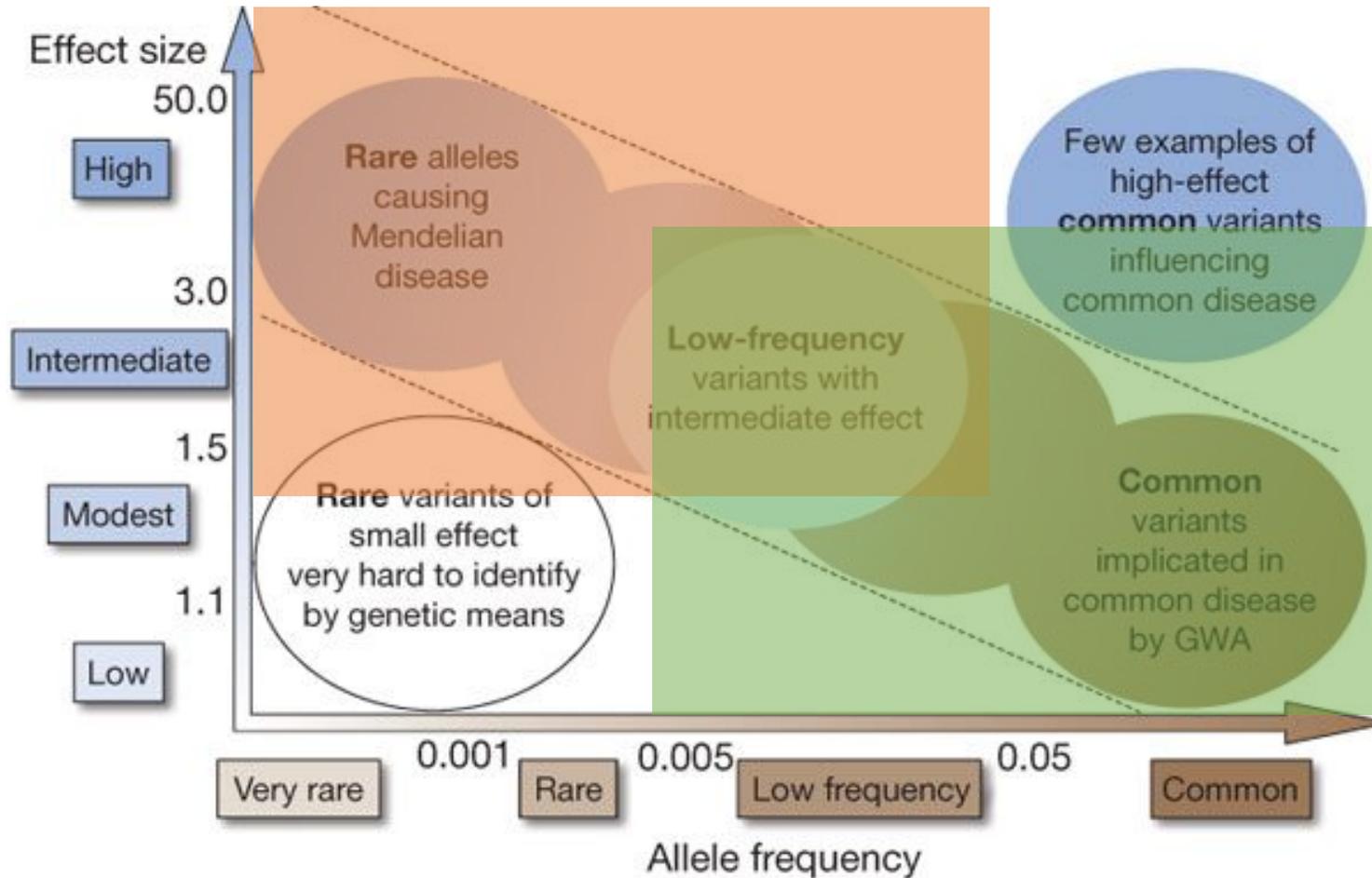
未診断患者数：推定37,000+名
(Orphanet J Rare Dis. 2018;13:208.)

- ・医療的不安
- ・経済的不安
- ・心理的不安



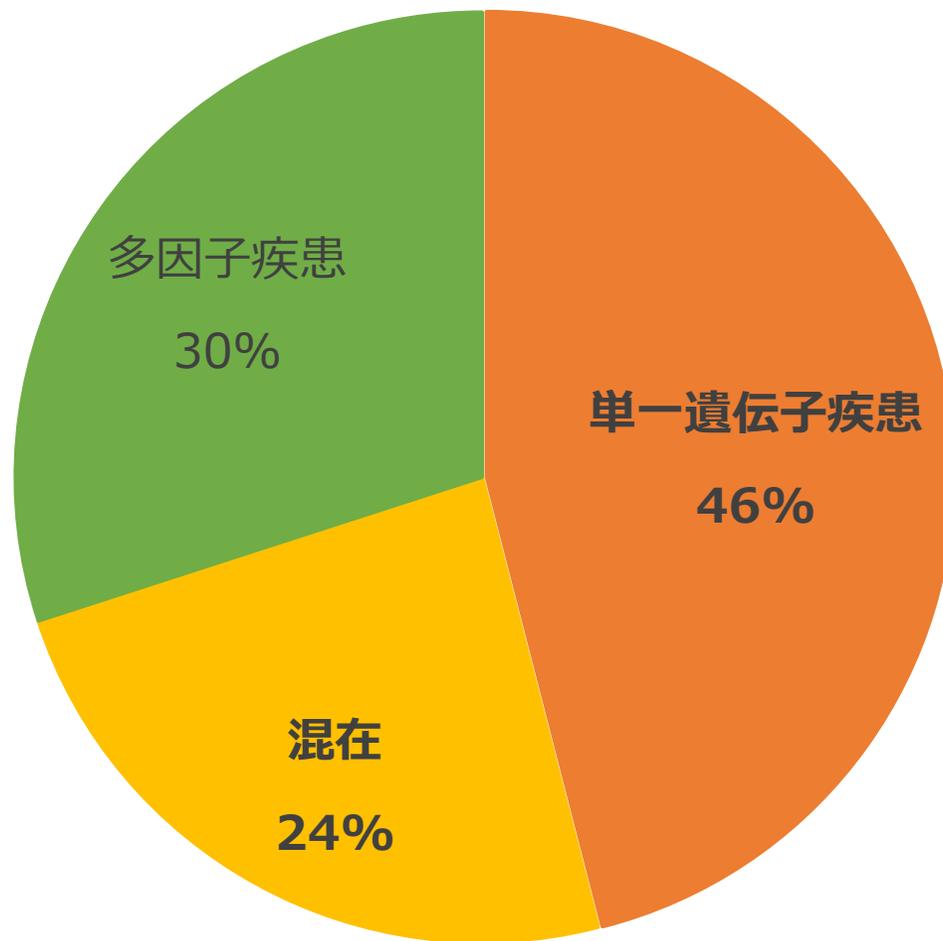
希少性・難治性疾患の病的バリエーション探索の背景

頻度の低いバリエーションほど、その効果は大きい ⇒ 単一遺伝子疾患

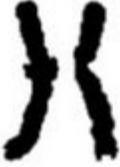
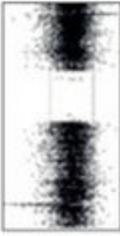


Manolio TA, et al. 2009. Nature 461(7265):747-753.

IRUD対象疾患：指定難病の例



ゲノム解析による診断

a	Light microscope	G-banded karyotype	Microarray	Whole-exome sequence	Whole-genome sequence
Appearance				<pre>CGGATGATTACCCGTT G.....GCTC TAGCTAGCTATA....</pre>	<pre>CGGATGATTACCCGTT GATATAGCTCTCGCTC GCTCTAGCTAGCTATA GGCTATGGGTGGGGGC</pre>
Resolution	Entire chromosome	5–10 Mb	50–100 kb	1 bp	1 bp
Number of loci probed	N/A	~500	~0.05–2 million	~50 million	3 billion
Variants detected	Aneuploidy, polyploidy	Variants >5 Mb	Copy number variants	Coding regions	Majority of variants
Variants per person	0 or 1	0 or 1	10–100s	~20,000	4–5 million
Diagnostic yield	Low				High
Incidental findings	Low				High

診断率
(一次の所見)

偶発的所見
(二次の所見)

Nat Rev Genet. 2018 May;19(5):253-268.

全エクソーム解析

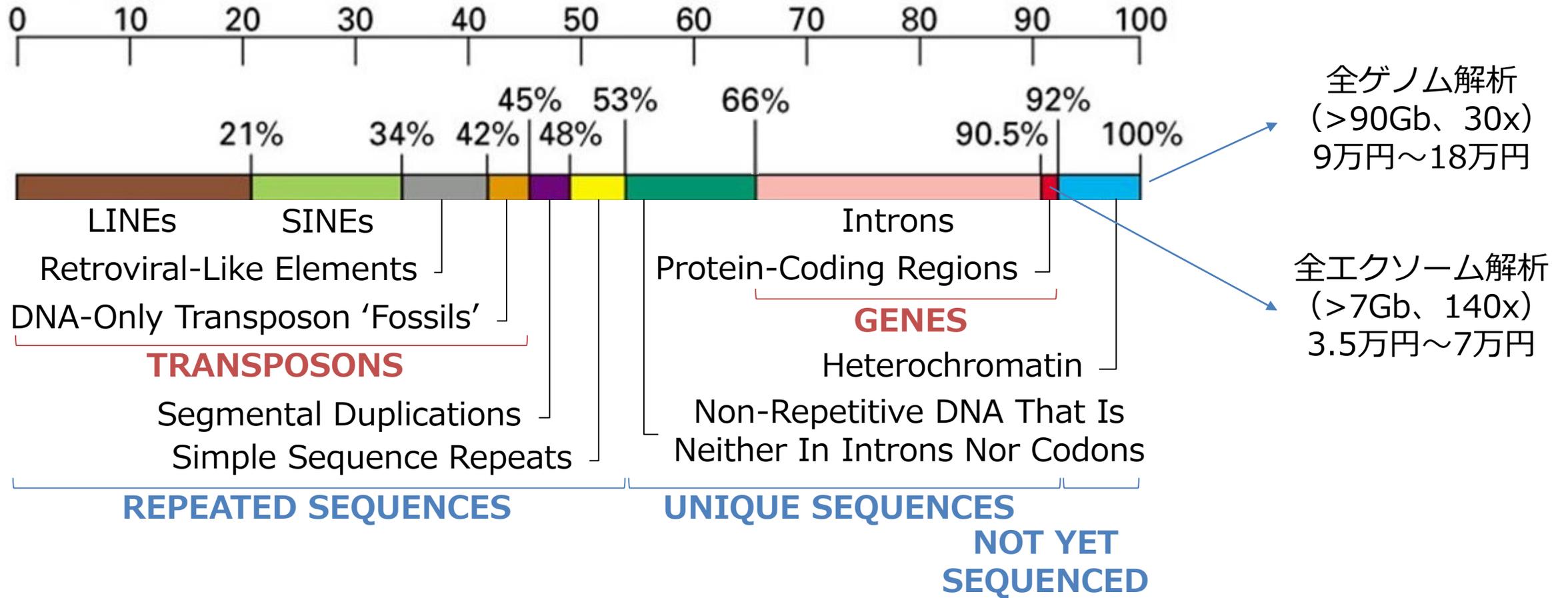
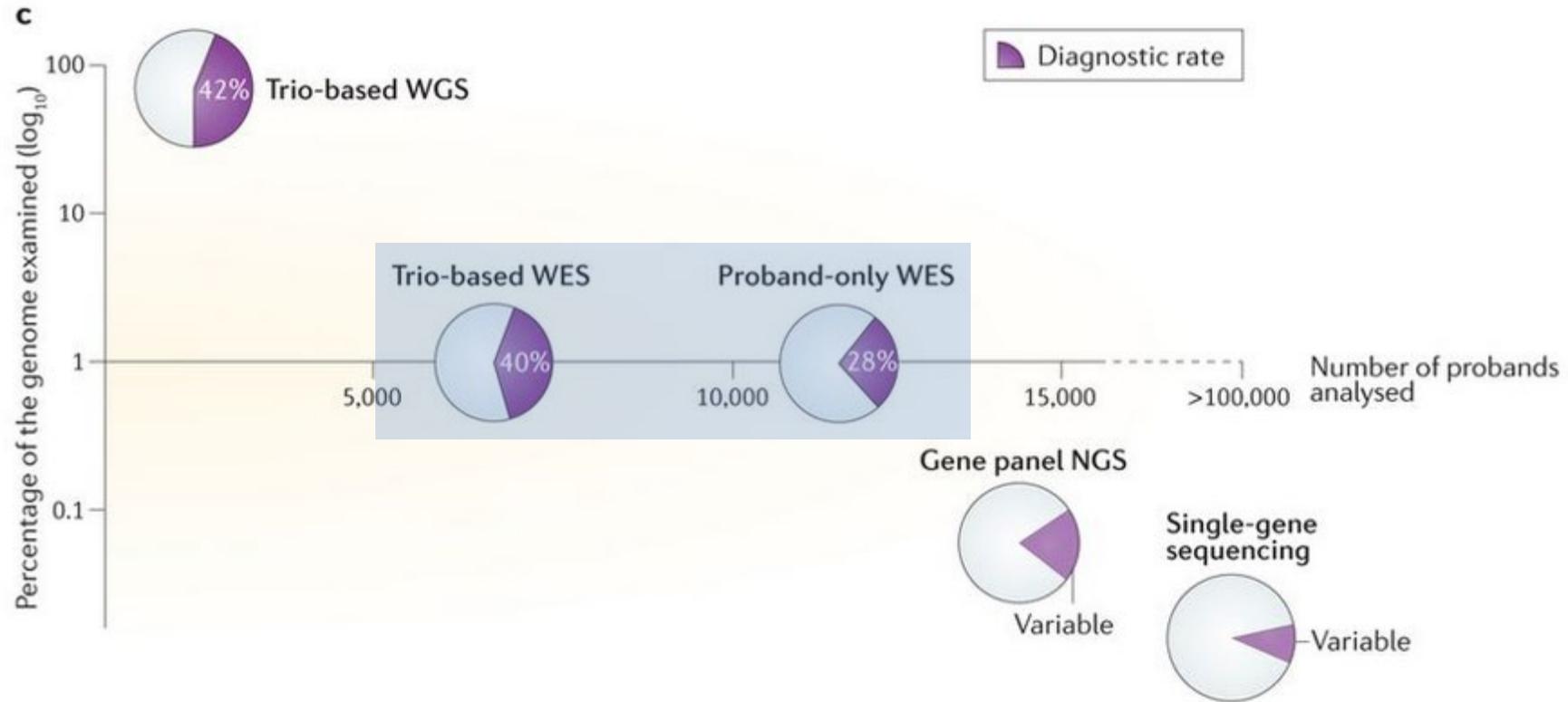


Figure 4-17. Molecular Biology of the Cell, 4th Edition.

診断率

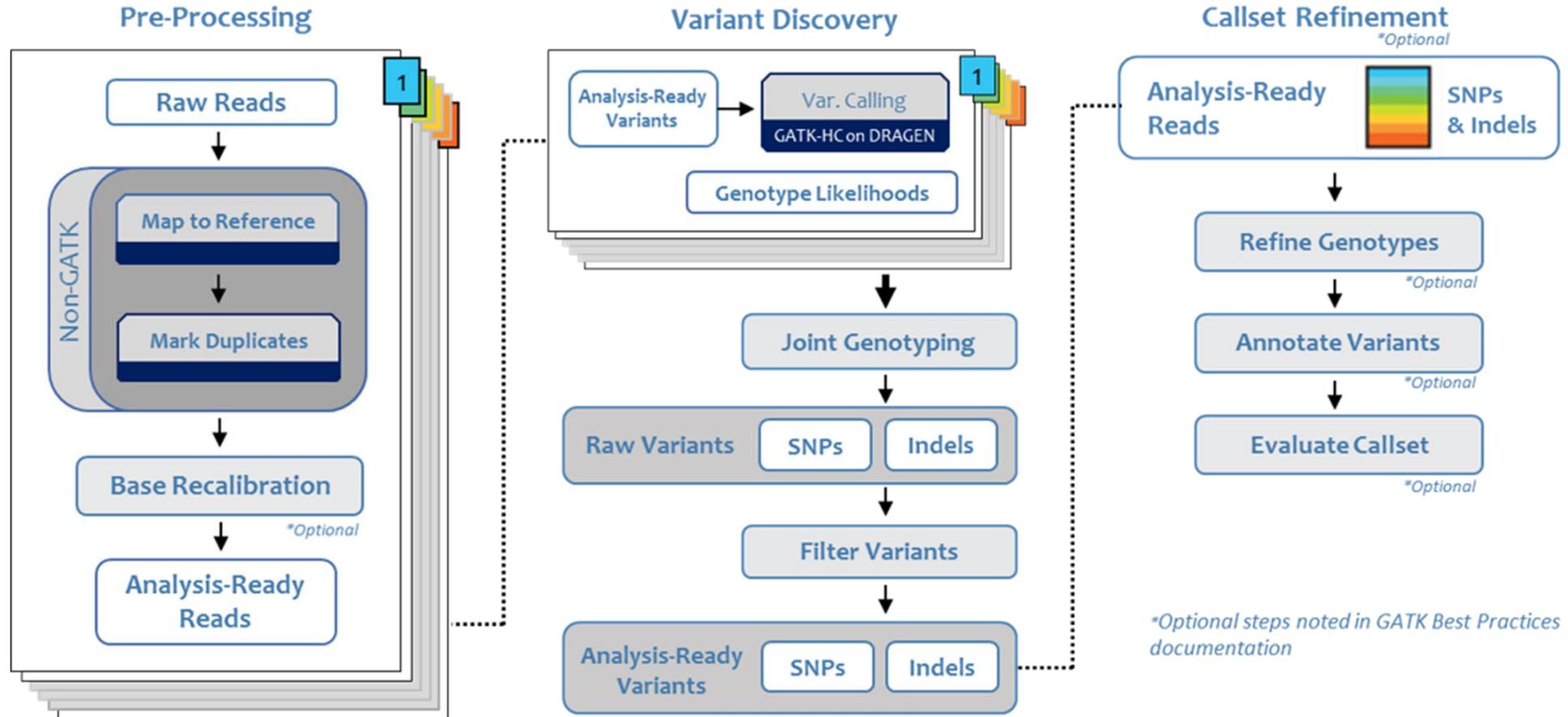


Nat Rev Genet. 2018 May;19(5):253-268.

次世代シーケンサー

プラットフォーム	NovaSeq 6000	HiSeq 2500	NextSeq 550	MiSeq
大規模全ゲノム	○	○		
大規模全エクソーム	○	○		
小規模全ゲノム			○	
小規模全エクソーム			○	
単一遺伝子				○
最大データ量	6,000 Gb	1,000 Gb	120 Gb	15 Gb
最大リード数	200億本	40億本	4億本	2,500万本
最大リード長	250 b x 2	250 b x 2	150 b x 2	300 b x 2

解析パイプライン



GATK Best Practices for Germline SNPs and Indels in Whole Genomes and Exomes on the DRAGEN Platform

マッピング



バリエーションコーリング



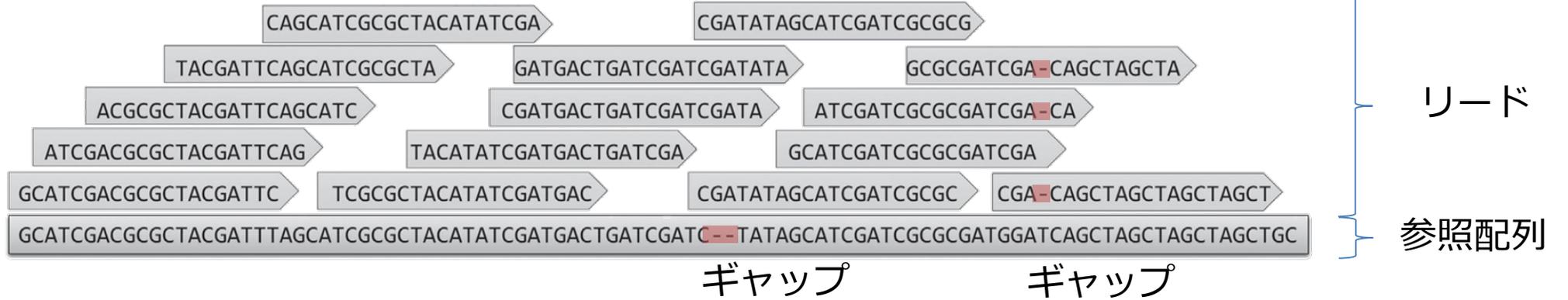
NextSeqオンラインセミナー

アノテーション

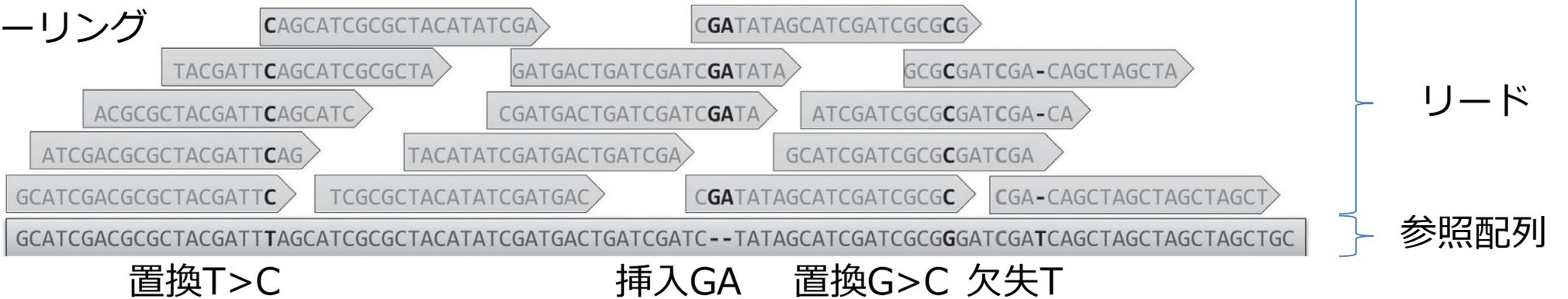


解析パイプライン (2)

① マッピング



② バリエントコーリング

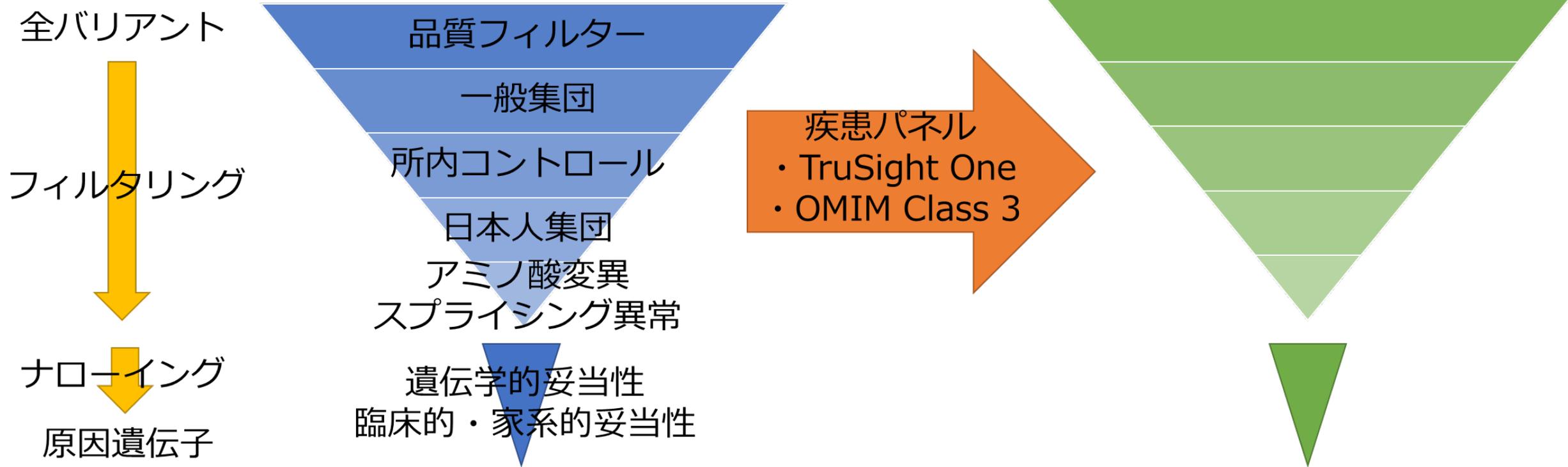


③ アノテーション

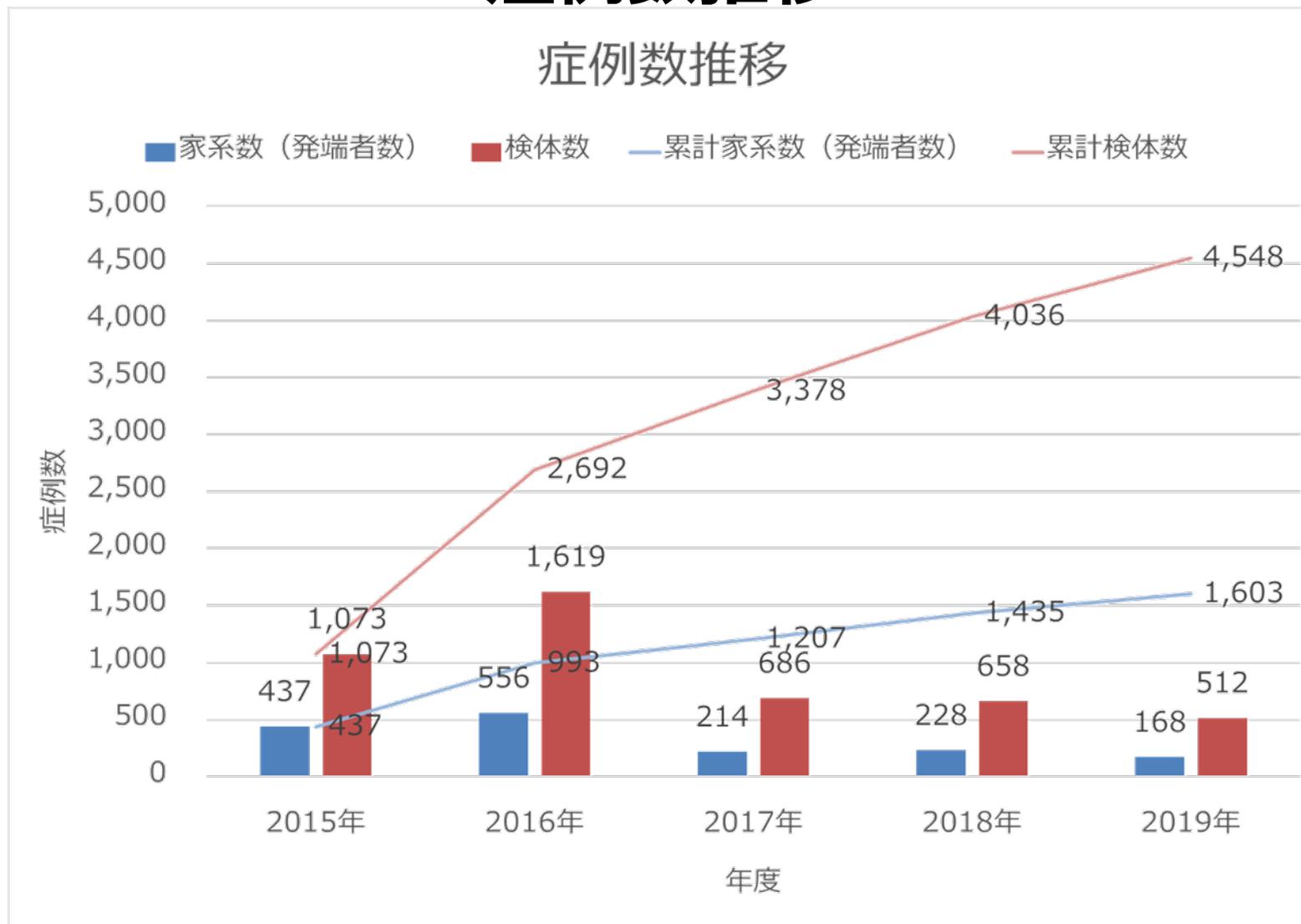
タンパク質構造・機能の変化あり? ⇒ (既知/新規) 病因遺伝子/バリエント?



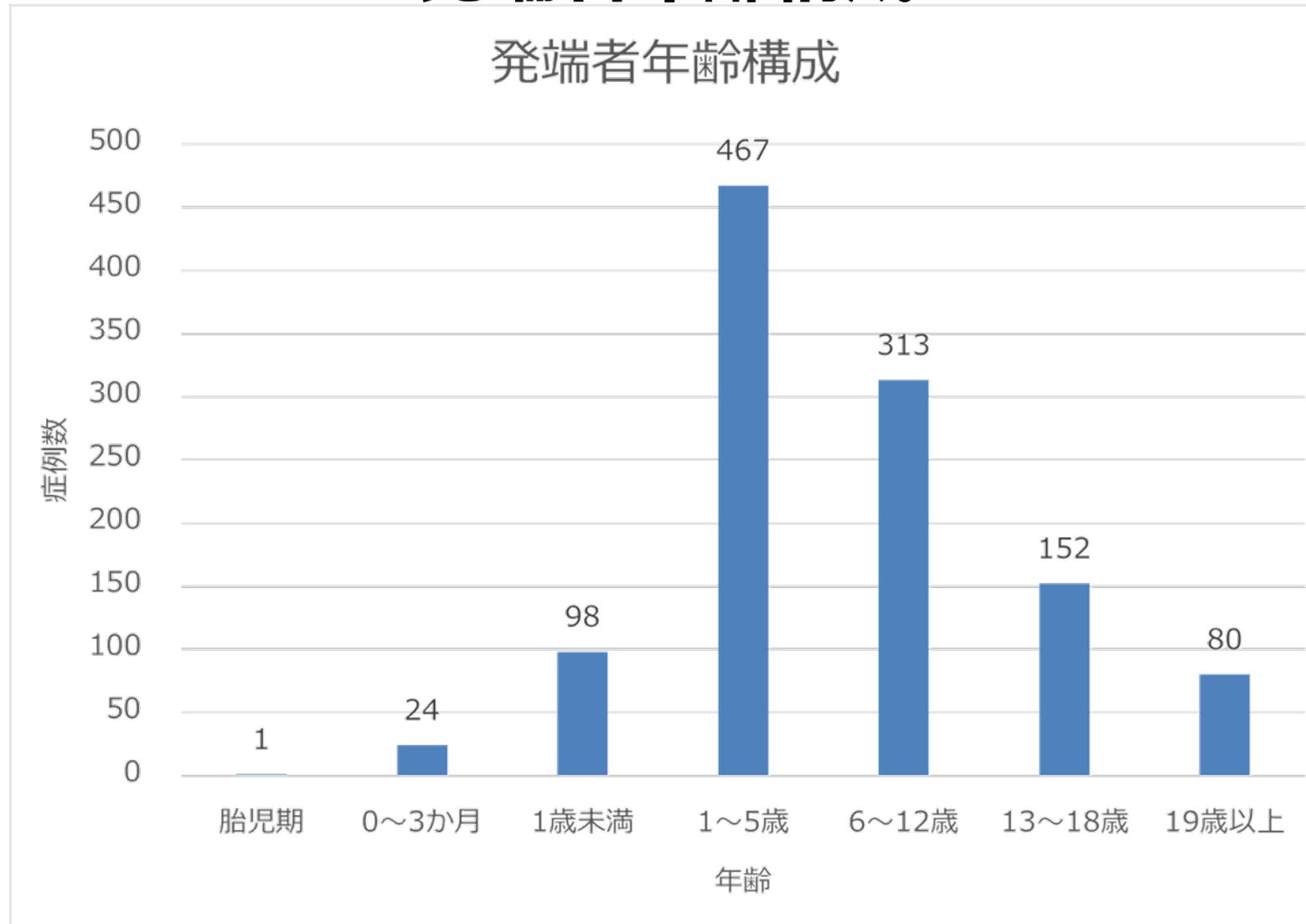
原因遺伝子の同定



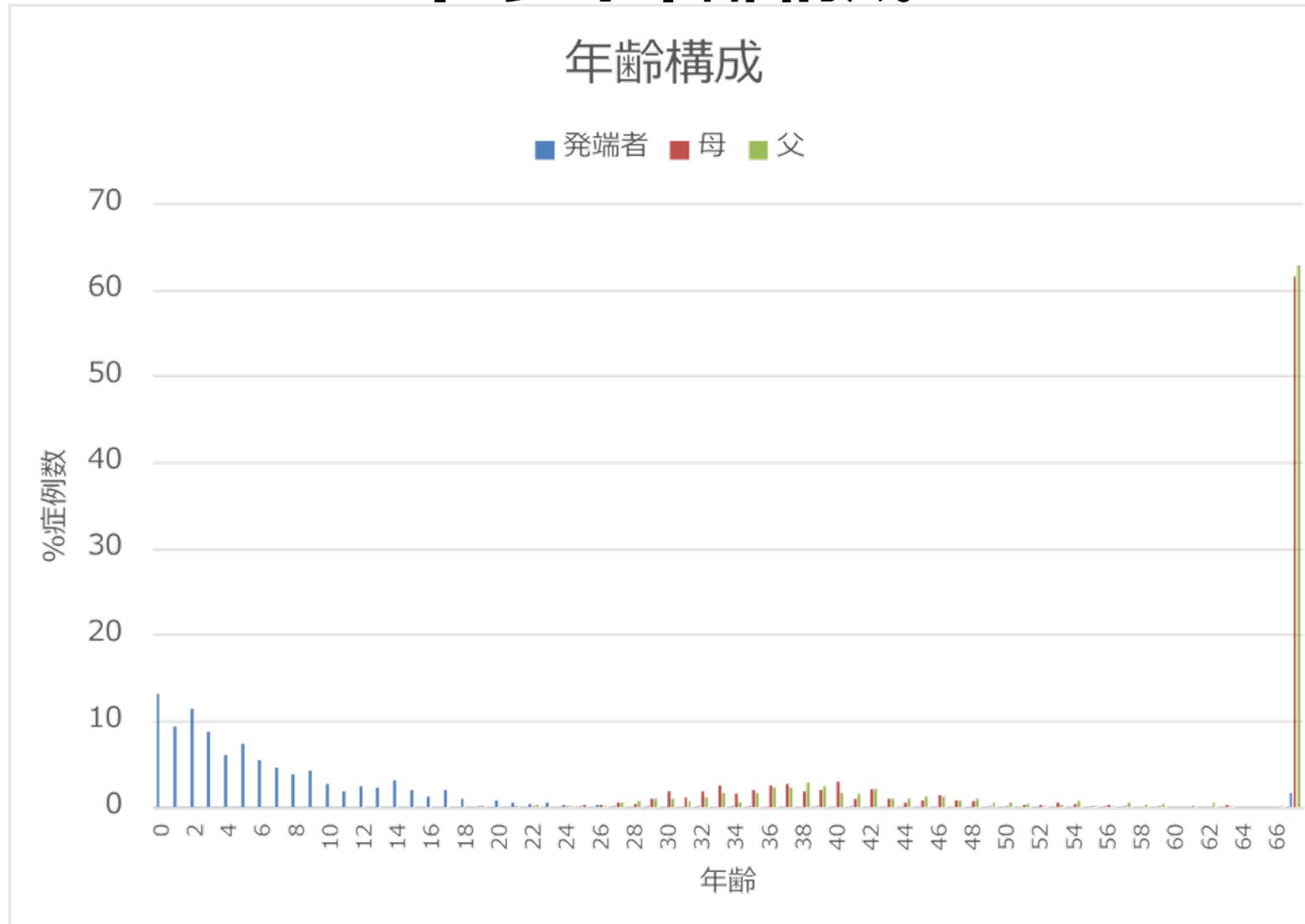
症例数推移



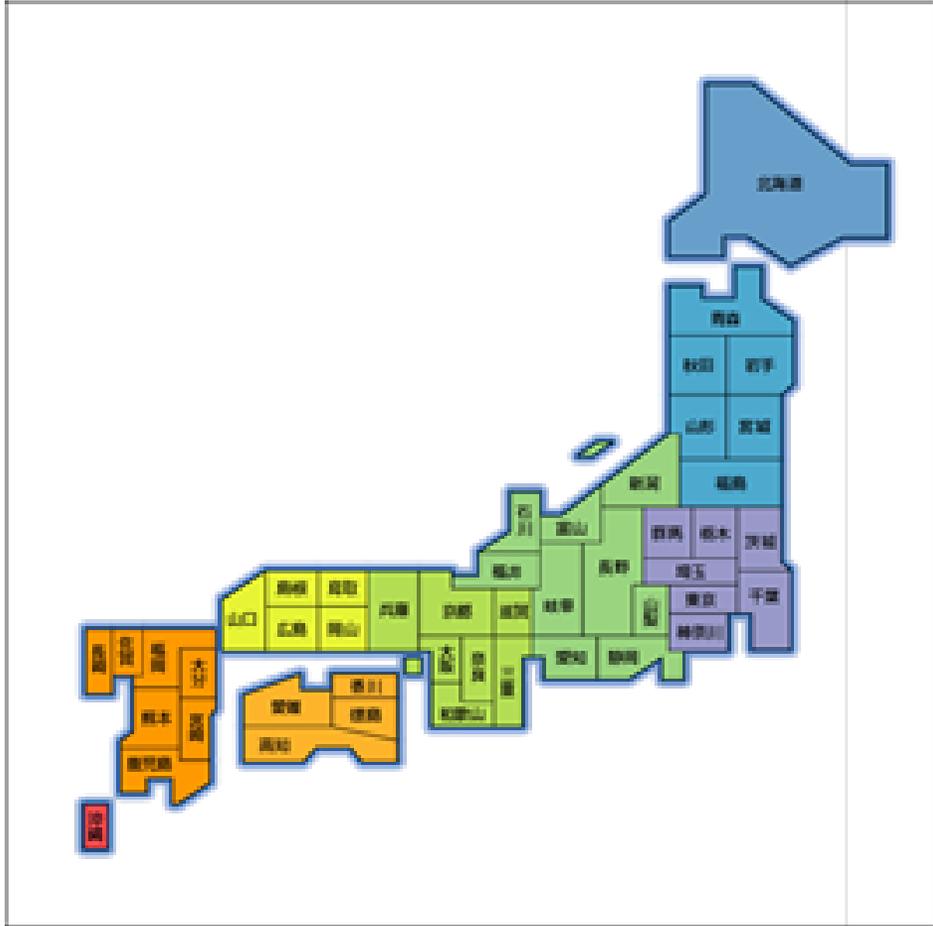
発端者年齢構成



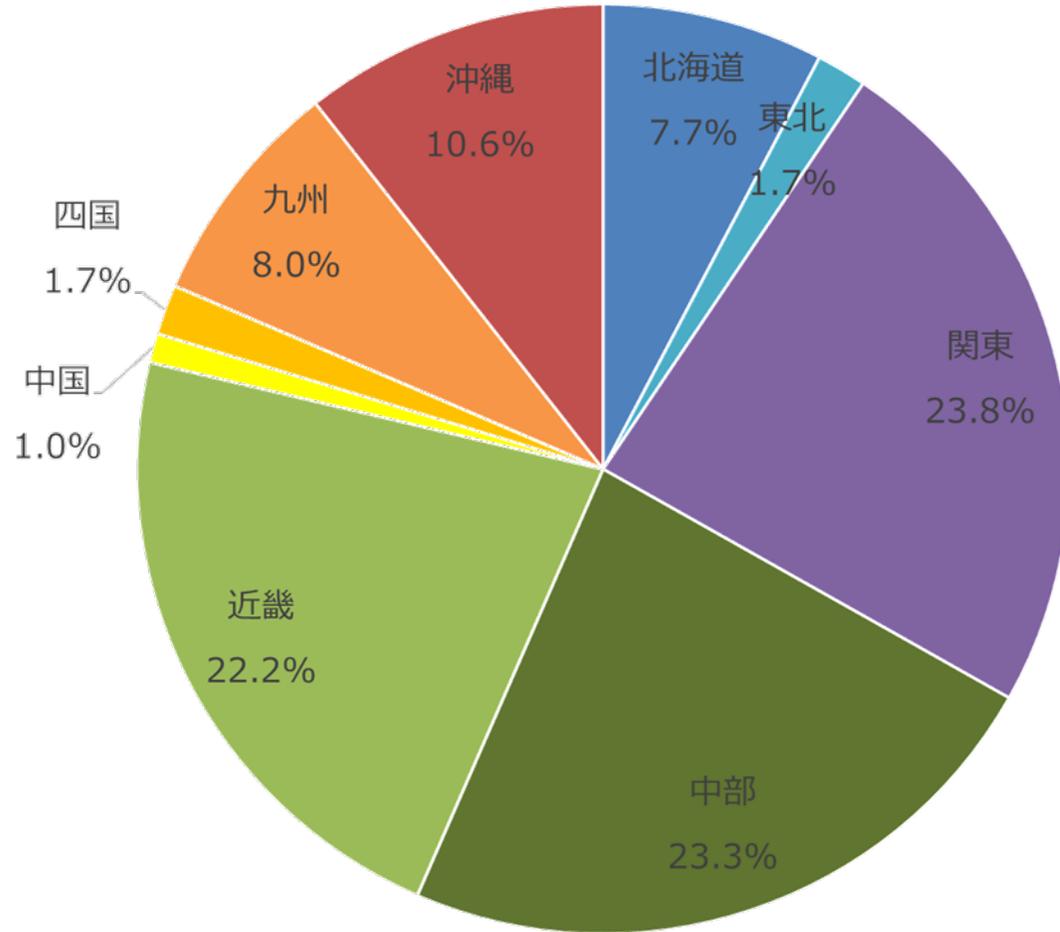
トリオ年齢構成



症例数地方別内訳

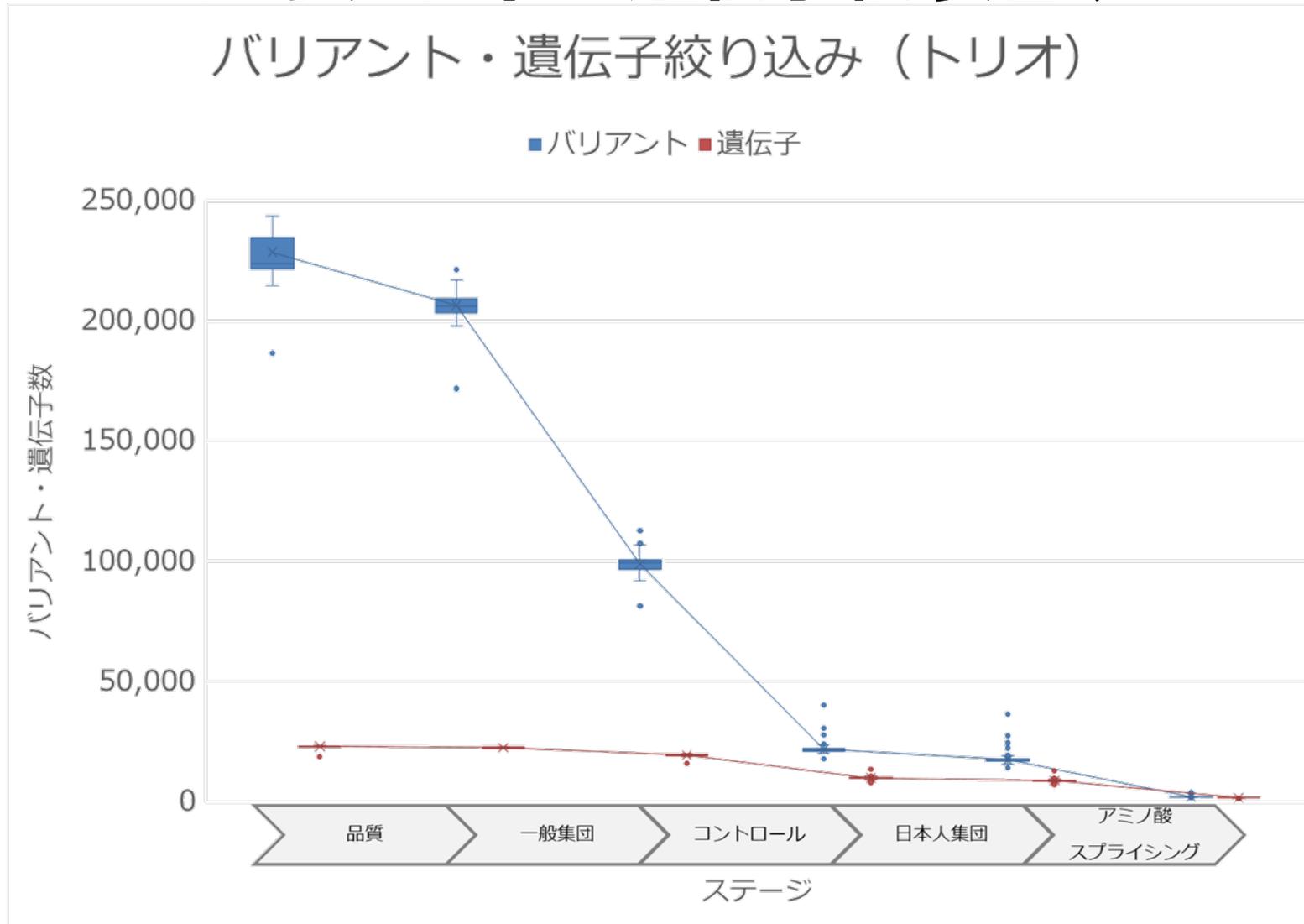


症例数地方別内訳



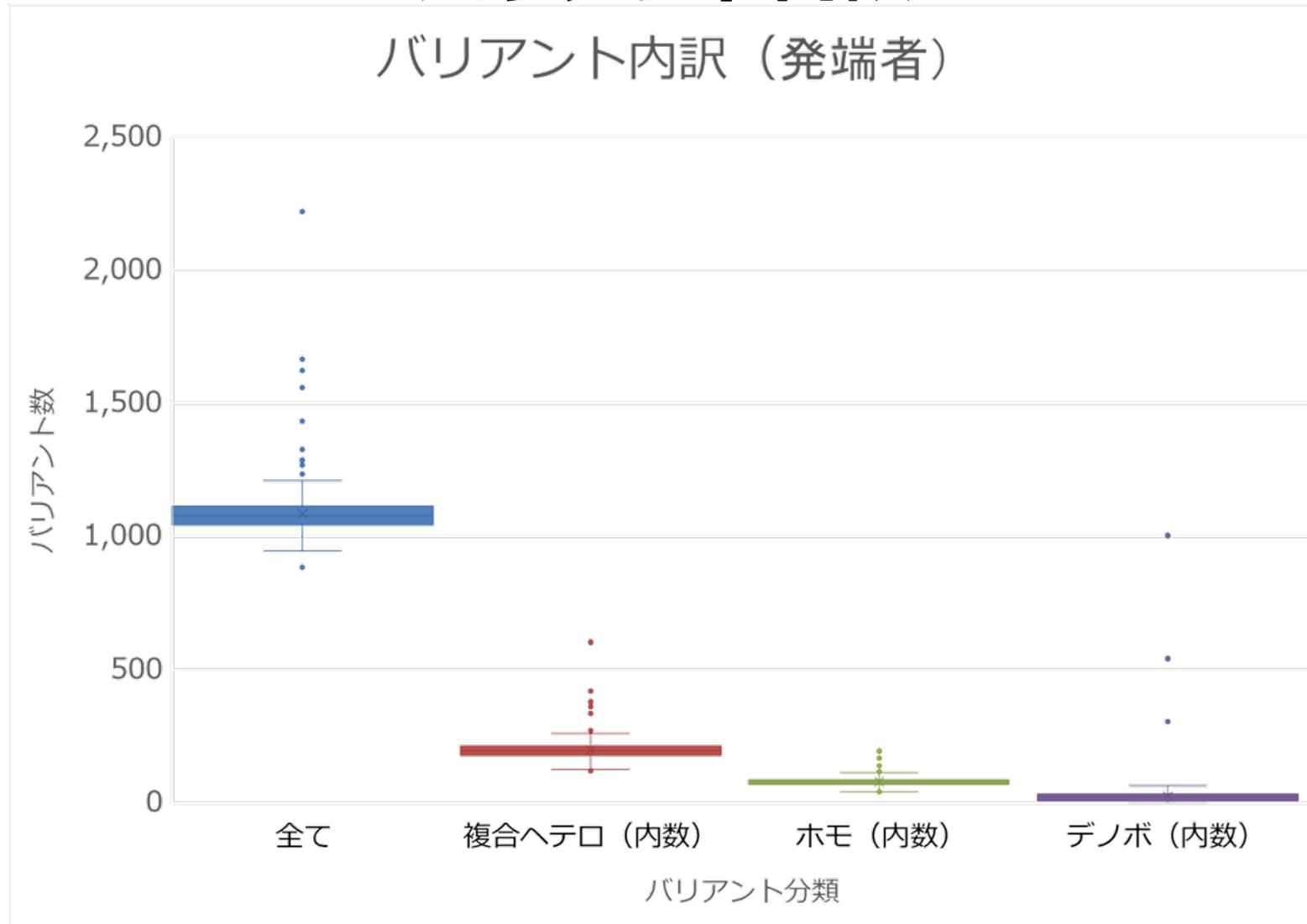
バリエーション・遺伝子絞り込み

バリエーション・遺伝子絞り込み（トリオ）



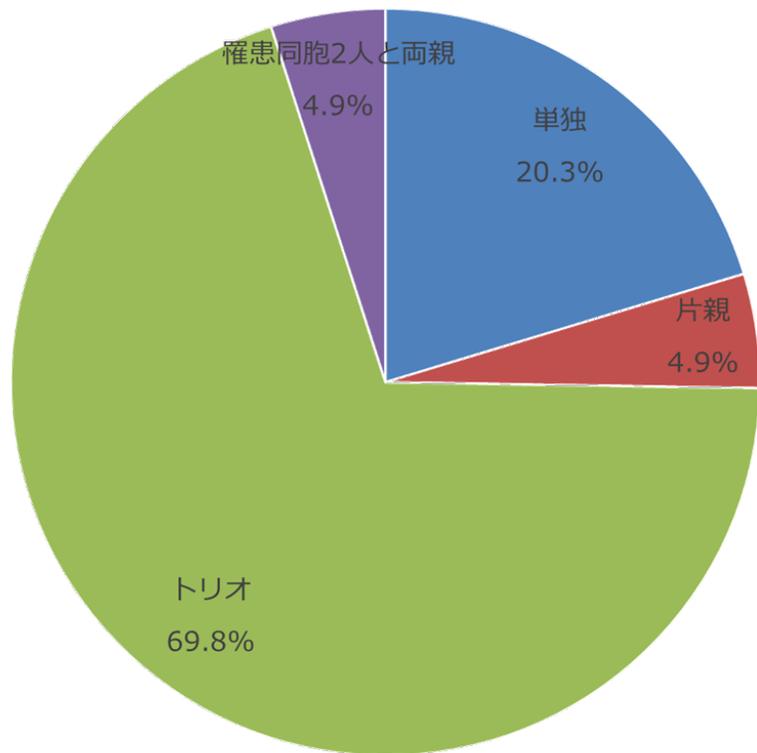
バリエーション内訳

バリエーション内訳（発端者）

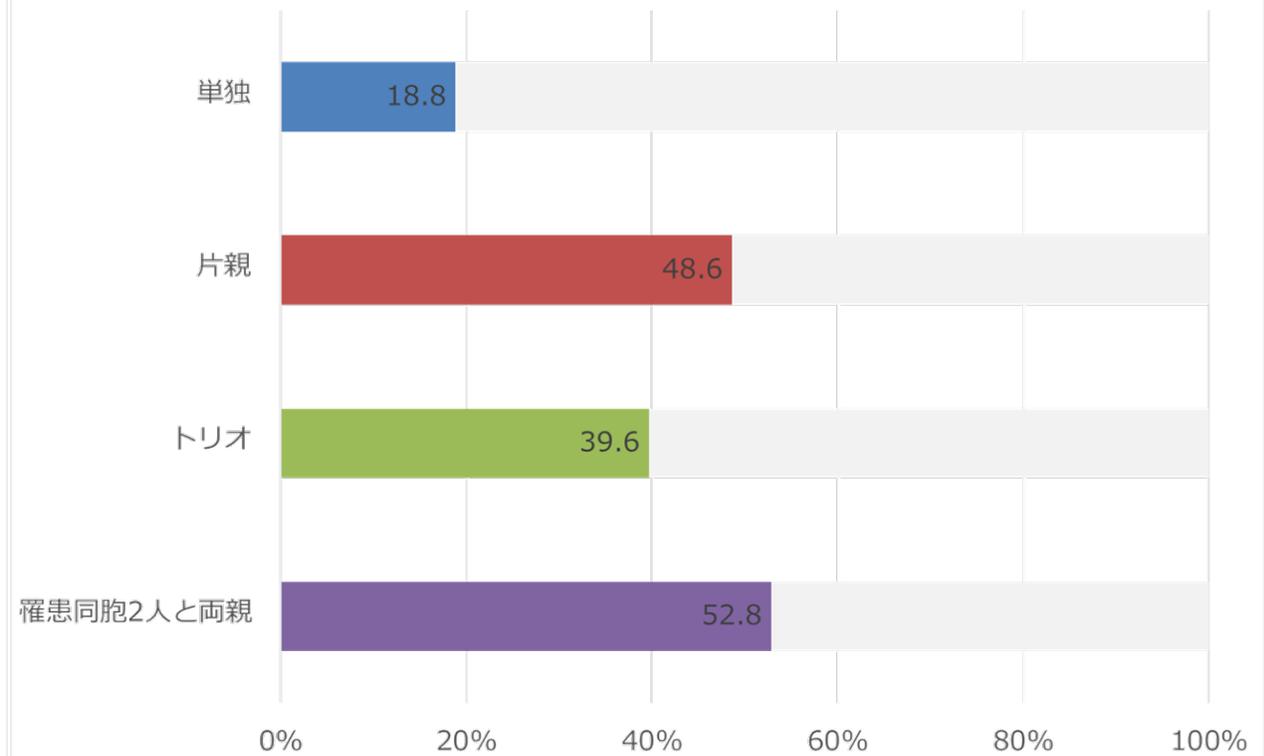


解析方法内訳・同定率

解析方法内訳

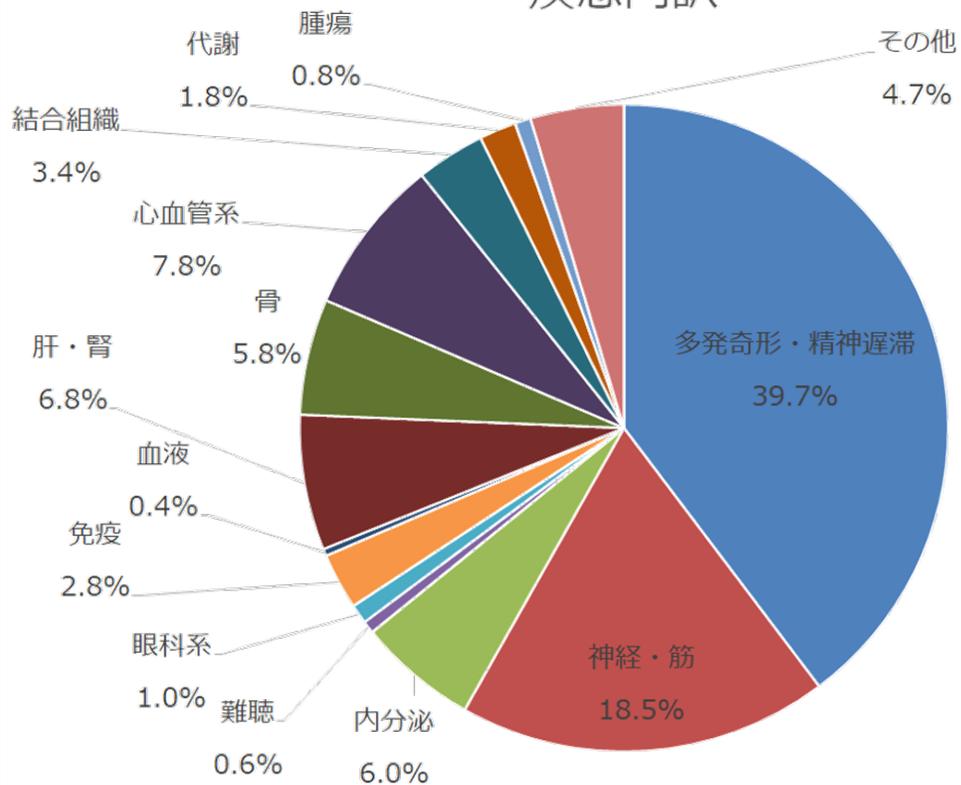


解析方法別同定率

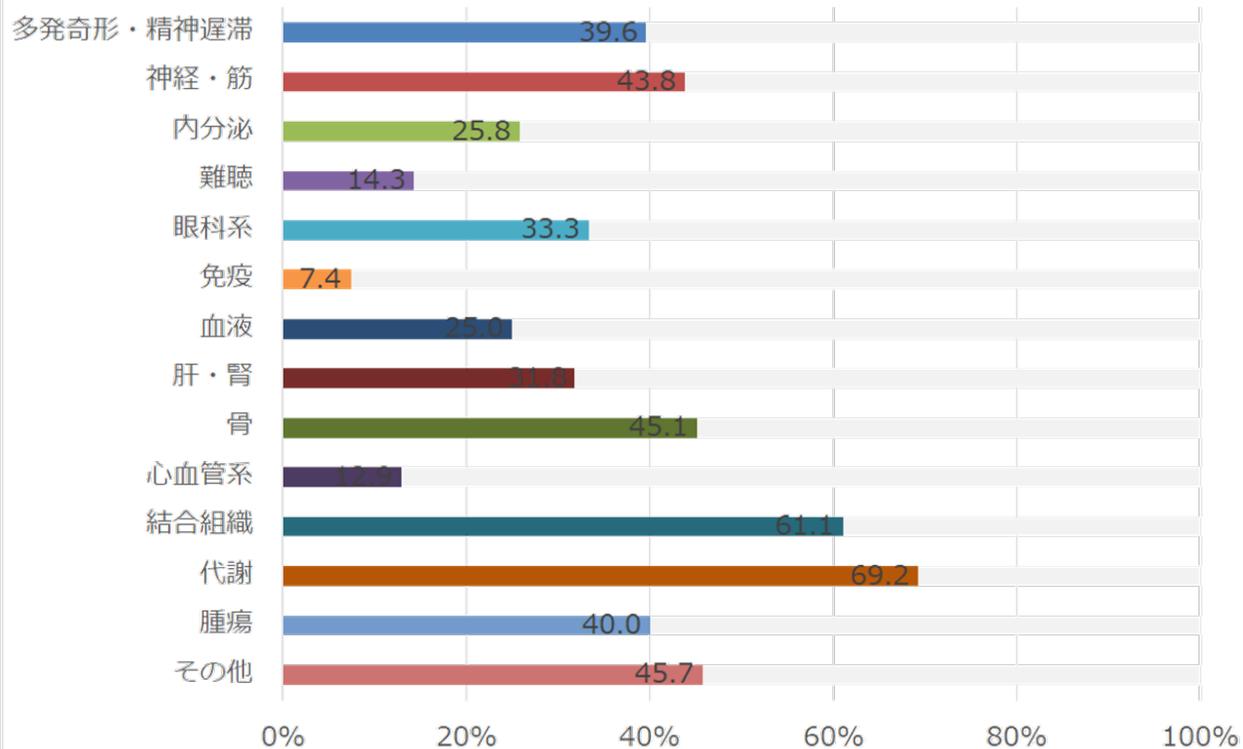


疾患内訳・同定率

疾患内訳

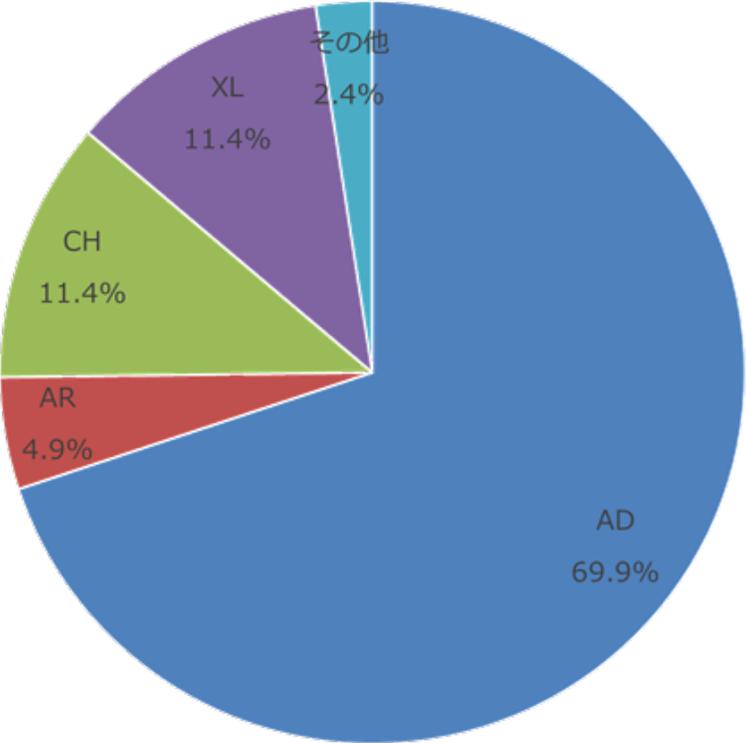


疾患別同定率

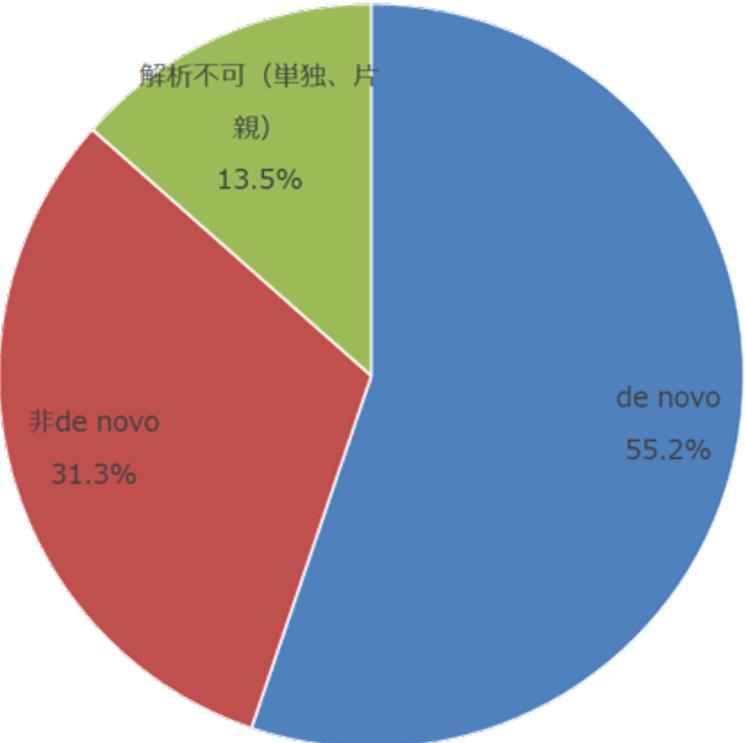


遺伝様式比率およびde novo比率

遺伝様式比率

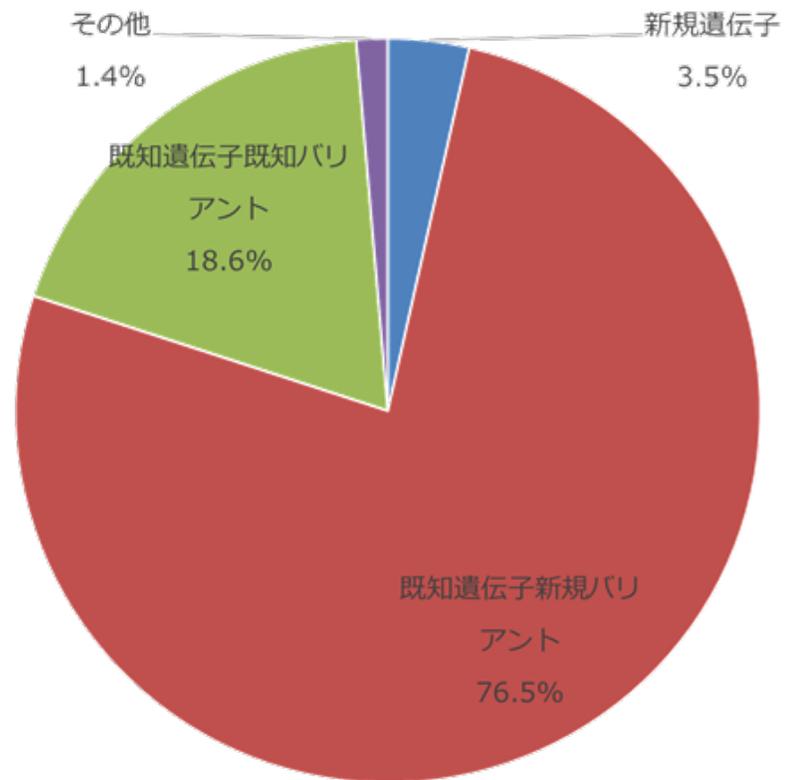


de novo比率

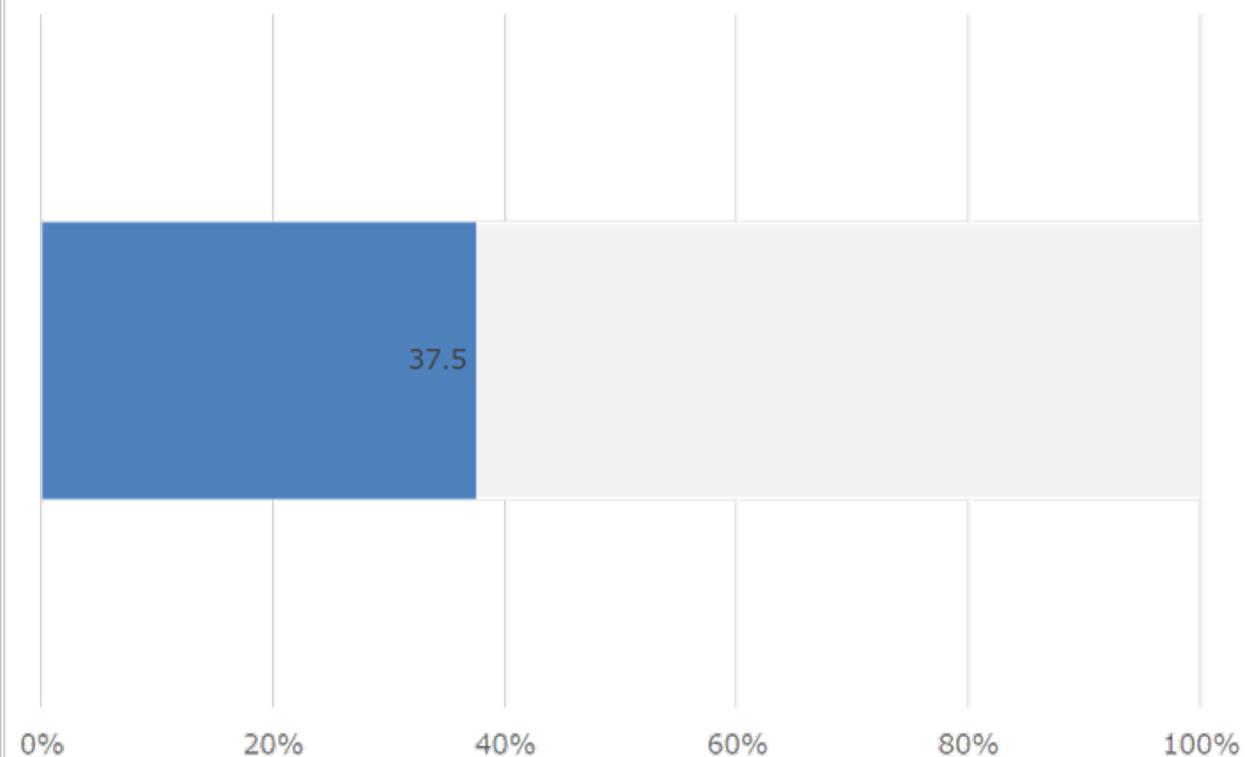


バリエーション分類および同定率

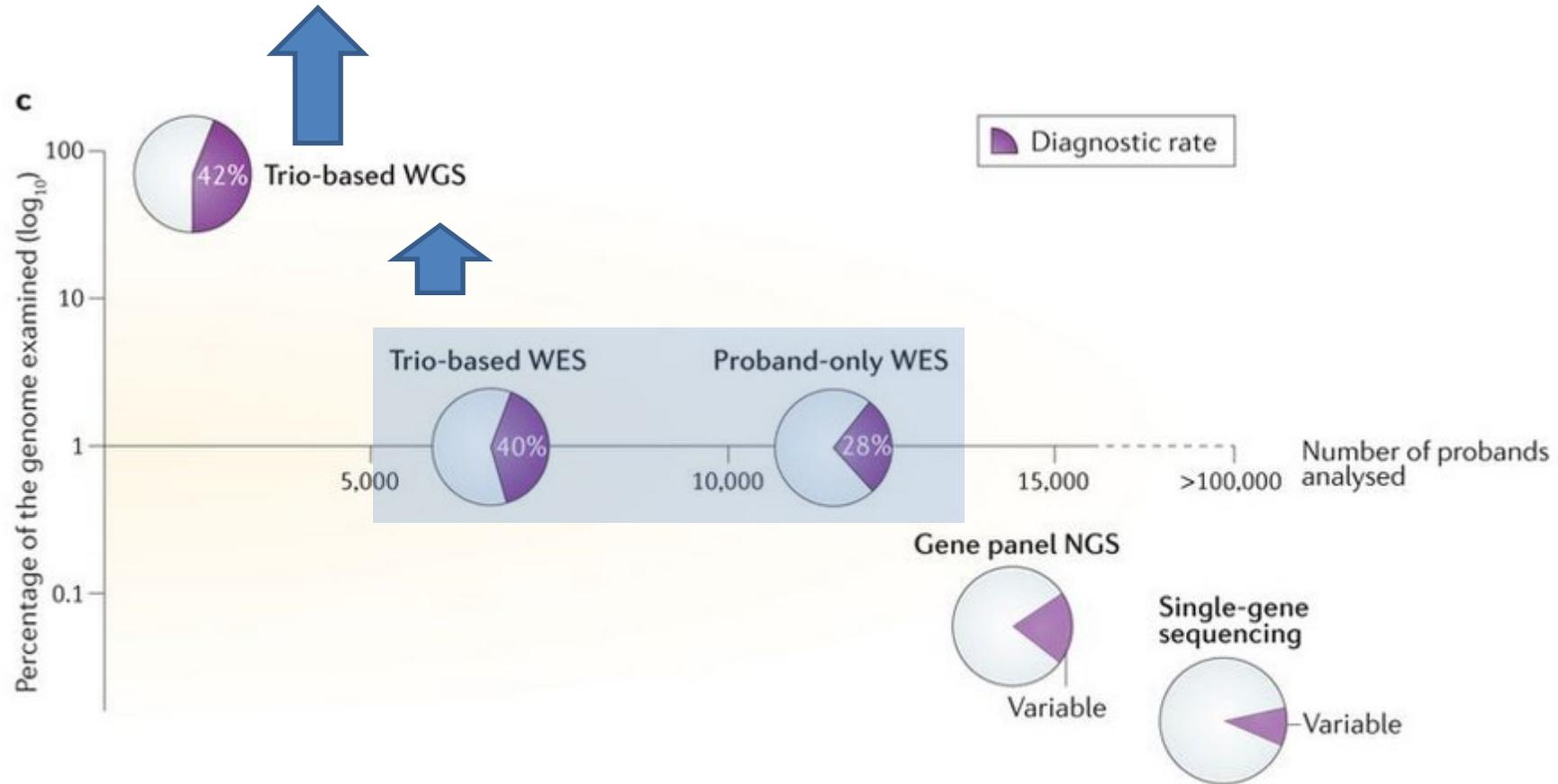
バリエーション分類



同定率



診断率の向上へ向けて



Nat Rev Genet. 2018 May;19(5):253-268.

WESからWGS、その先

がん種	必要な検体の種類	先行解析		本格解析
		解析対象の最大症例数	先行解析の症例数	
罹患数の多いがん・難治性がん	がん部位（新鮮凍結検体） + 血液（正常検体）	約5.6万症例（11.2万ゲノム）+α	難治性のがん、希少がん、遺伝性のがんの約1.6万症例（3.3万ゲノム）+β	先行解析の結果や国内外の研究動向等を踏まえて検討
希少がん（小児がんを含む）		約0.7万症例（1.4万ゲノム）+α		
遺伝性のがん（小児がんを含む）		約0.2万症例（0.4万ゲノム）+α （必要に応じて両親や同胞の正常検体0.19万ゲノム）		
合計		約6.4万症例（約13万ゲノム）+α	-	-

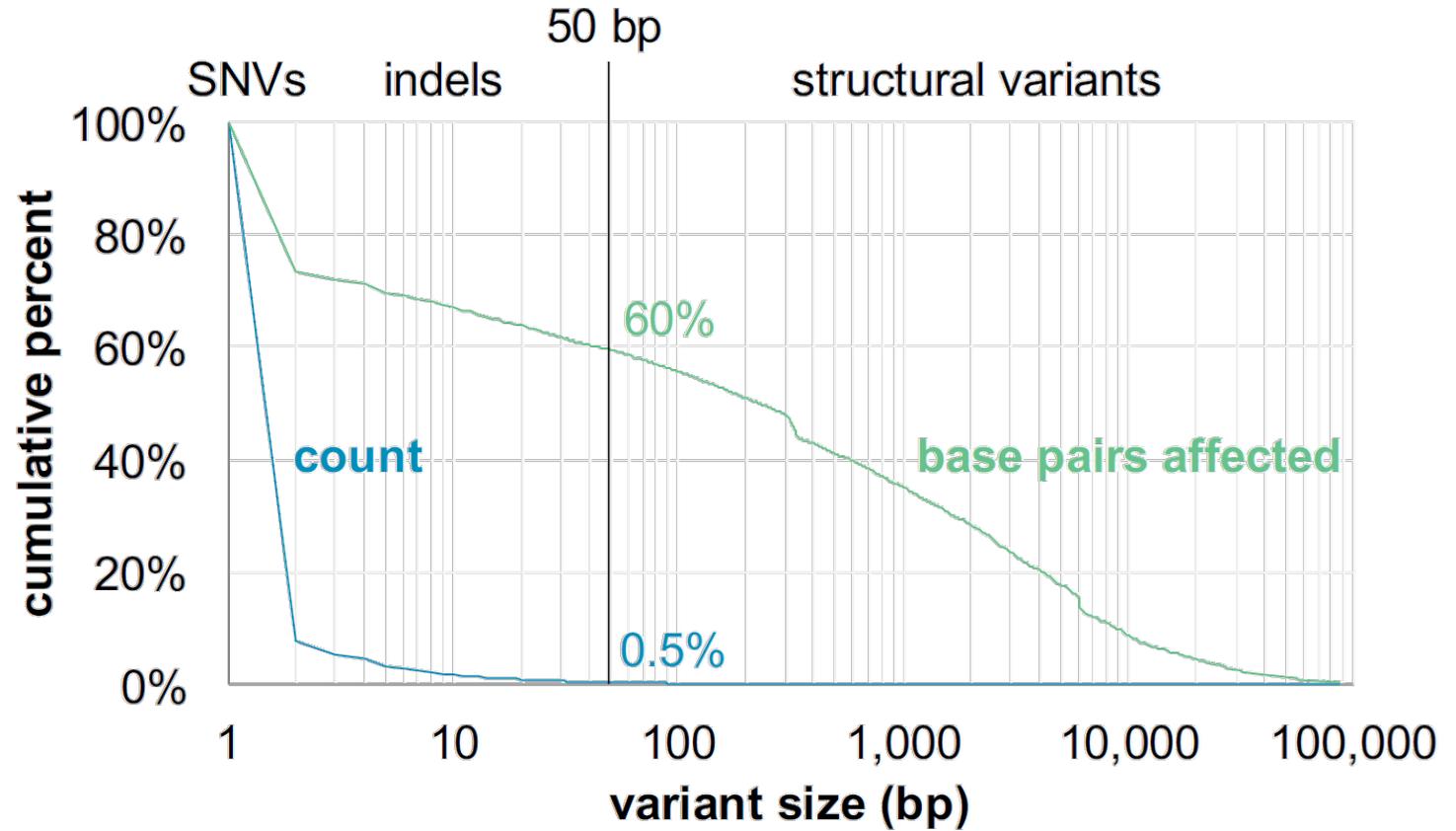
対象類型	必要な検体の種類	先行解析		本格解析
		解析対象の最大症例数	先行解析の症例数	
単一遺伝子性疾患（筋ジストロフィー等）	血液（加えて両親の血液）	約0.1万症例（0.3万ゲノム）+α	単一遺伝子性疾患、多因子疾患、診断困難な疾患の約5500症例（6500ゲノム）+β	先行解析の結果や国内外の研究動向等を踏まえて検討
多因子疾患（パーキンソン病等）	血液	約2.4万症例（2.4万ゲノム）+α		
診断困難な疾患	血液（加えて両親の血液）	約0.3万症例（0.9万ゲノム）+α		
合計		約2.8万症例（約3.6万ゲノム）+α	-	-

厚生労働省 全ゲノム解析等実行計画（第1版）より

ヒトゲノムにおける構造多型

構造多型	疾患例
挿入 	Charcot-Marie Tooth疾患, Tay-Sachs疾患
欠失 	Williams症候群, Duchenne筋ジストロフィー, Smith-Magenis症候群, Carney Complex
重複 	アルツハイマー病におけるAPP, Potocki-Lupski症候群, Prader-Willi症候群, Angelman症候群
転座 	ダウン症候群, XX男性症候群 (SRY), 統合失調症 (chr 11), パーキットリンパ腫
逆位 	Hemophilia A, Hunter症候群, Emery-Dreifuss筋ジストロフィー
縦列反復 	脆弱X症候群におけるFMR1, ハンチントン病, 脊髄小脳失調

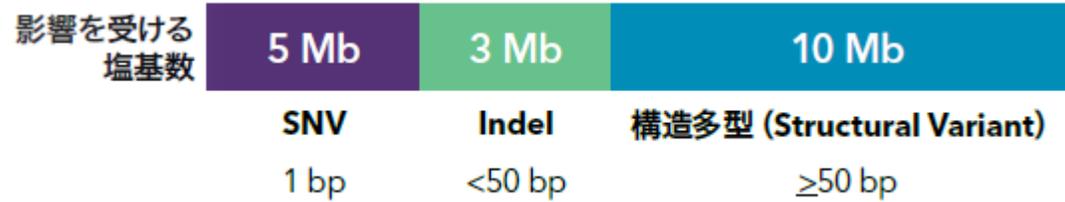
Pacific Biosciences社のパンフレットより



Wenger A, et al. ASHG 2017.

Only 0.5% of variants but 60% of variant base pairs are in structural variants ≥ 50 b.

診断率向上への挑戦



Pacific Biosciences社のパンフレットより

ショートリードシーケンサー

読み取り長：~100b
読み取り精度：~99%



ロングリードシーケンサー



読み取り長：~50,000b
読み取り精度：~90%

ゲノムマッピングシステム



ある難聴疾患原因遺伝子のアレル頻度の地域差

	全県	沖縄県	沖縄県以外
患儿	0.0061	0.0339	0.0023
家族	0.0062	0.0289	0.0032
全員	0.0061	0.0306	0.0029

考察

- これまでに1,603家系（4,548検体）のWESを行った。
- 全体での同定率は37.5%であったが、疾患や解析方法によって異なった。
- 同定した遺伝子の3.5%は新規遺伝子、76.5%は既知遺伝子新規バリエーション、18.6%は既知遺伝子既知バリエーションであった。
- アノテーションDBの選択、候補バリエーションのフィルタリングやナローイングの方針は重要である。
- 同定率の向上のためには、ロングリードシーケンサー等の新技術の利用が不可欠である。
- 遺伝的プロファイルは地域集団によって大きく異なることがあり、注意が必要である。