

---

# **ゲノム・オミックス統合解析WG 中間とりまとめ 第2段階におけるゲノム・オミックス解析の実施計画**

---

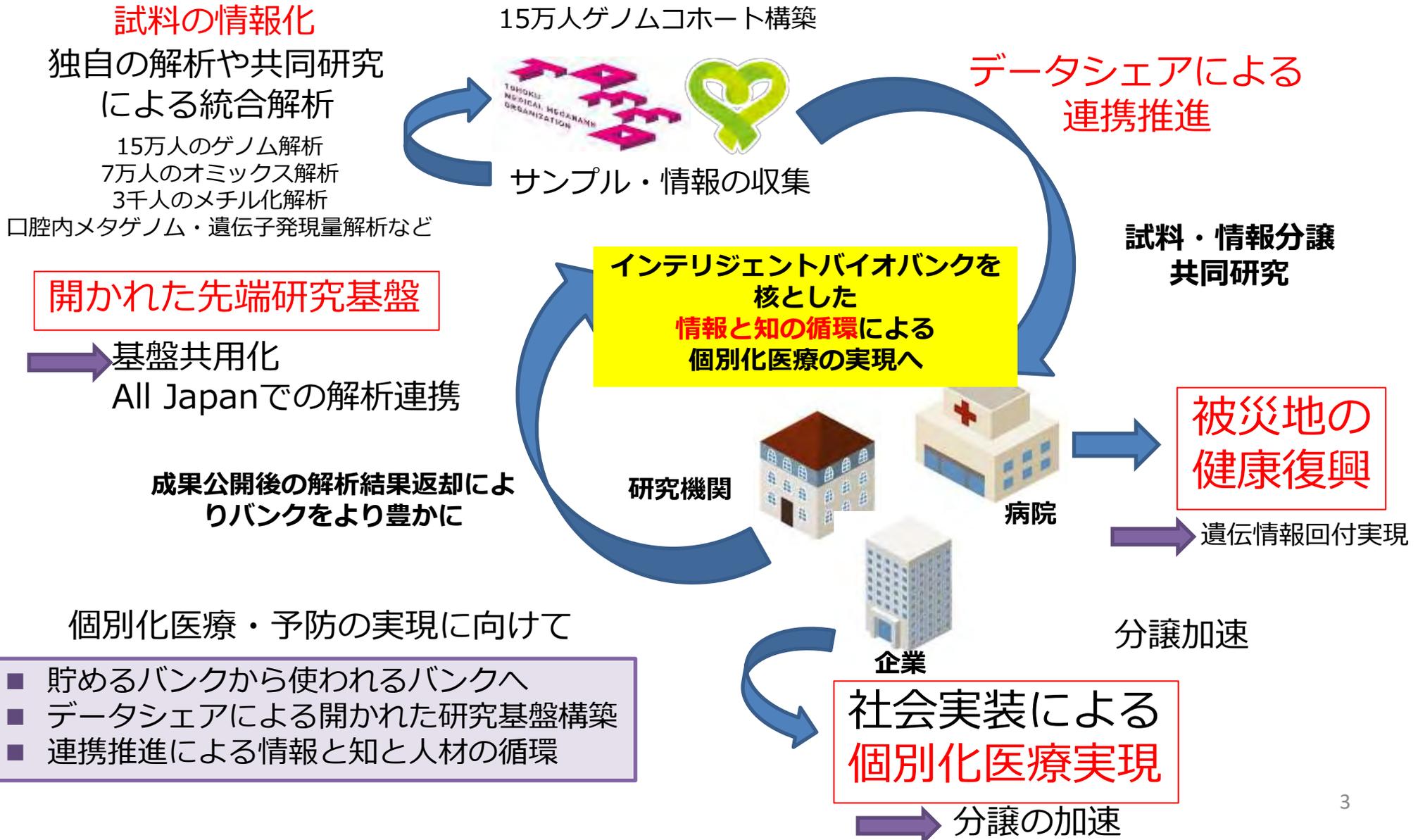
# ゲノム・オミックス解析

## 第2段階計画概要

---

1. はじめに
2. 大規模ゲノム解析の意義
  - 1) 海外のコホート・バイオバンクにおけるゲノム医学研究の現状、大規模アレイ解析との比較
  - 2) 疾患関連遺伝子研究に向けた基盤の提供
  - 3) 大規模ケース・コントロールスタディでの活用
3. ゲノム解析の実施計画
  - 1) ゲノム解析のロードマップ
  - 2) 大規模SNPアレイ解析と全ゲノム情報復元
  - 3) 全ゲノムリファレンスパネルの拡充
  - 4) 全ゲノム解析の手法とその選択
4. オミックス解析等の実施計画
  - 1) 多層オミックス参照パネルの構築と展開
  - 2) 大規模メタボローム解析
  - 3) エピゲノム解析
  - 4) その他のオミックス解析等
5. 個別化予防に向けた統合解析のための基盤構築
  - 1) 発症リスク予測手法開発
  - 2) 統合・知識データベースと病院連携
6. 遺伝情報回付
7. ゲノム医療実現のための人材育成
8. まとめ

# 実施計画案概要



# 第2段階計画の概要

## ゲノム医療研究の基盤構築

- 全参加者を対象として標準ゲノムレファレンスパネルを用いて開発したアレイによるゲノム解析を行う。また、多くの参加者を対象に疾患発症と関連するマーカー探索に向けたオミックス解析を実施する
- 第1段階と合わせて8千人の全ゲノム解析を行い、精度の高い日本人の標準ゲノムレファレンスパネルを構築する
- 早期の収集した試料・情報の分譲を行うとともに、他コホート事業やバイオバンクと連携しつつ、保管試料・情報のカタログ化や横断検索できるシステムの導入など、安全性を確保しつつユーザーの利便性向上に取り組む
- ゲノム医療実現推進プラットフォーム事業により委託されているスーパーコンピュータ等のゲノム解析基盤を共用できる体制を構築する

## 個別化医療・個別化予防の先導モデルの構築

- 遺伝情報回付パイロット研究を推進する
- 高血圧、アトピー性皮膚炎、脳梗塞等の疾患発症リスク予測手法を開発するとともに、個人毎に疾患発症リスクを回付するための手法の開発も併せて行う

## ゲノム医療実現のための環境整備等への貢献

- ゲノム医療体制の構築に必要な専門人材を育成するとともに、そのキャリアパスの形成を図る。また、メディカルクラーク取得者と同程度の技能を有するデータマネージャーを育成する
- ゲノム医療を社会に根付かせるための取組を実施するとともに、ゲノム医療に関する倫理的諸課題の解決に取り組み、ゲノム医療を実現する素地を醸成する

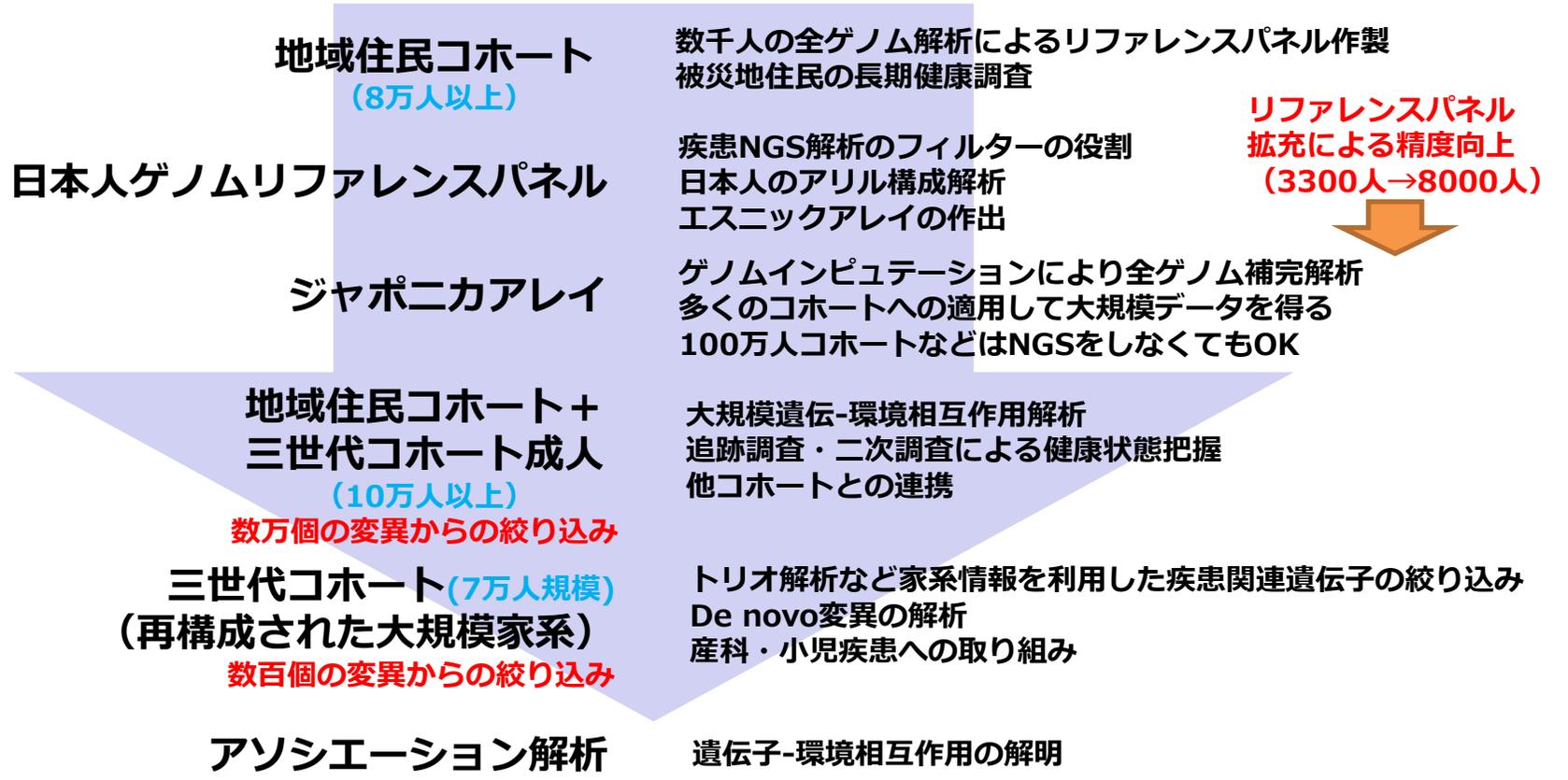
# ゲノム解析研究の戦略

アイランド  
deCODEジェネティクス

## TMMのコホートデザインの特徴

地域住民コホートを基盤とした全ゲノム解読とそれに基づくエスニックアレイ作出、同アレイを用いた全ゲノム解析に三世代コホートを用いた家系解析を組み合わせ、疾患関連遺伝子の特定と検証を目指す先進モデルである

deCODEの特徴  
全ゲノム解読とそれに基づくアレイ解析に広範な家系図を用いた解析を組み合わせ、次々と疾患原因遺伝子を特定している先進モデルであるが、企業が実施している点での限界もある



# ゲノム・オミックス解析

## 第2段階計画概要

---

### 1. はじめに

### 2. 大規模ゲノム解析の意義

- 1) 海外のコホート・バイオバンクにおけるゲノム医学研究の現状、大規模アレイ解析との比較
- 2) 疾患関連遺伝子研究に向けた基盤の提供
- 3) 大規模ケース・コントロールスタディでの活用

### 3. ゲノム解析の実施計画

- 1) ゲノム解析のロードマップ
- 2) 大規模SNPアレイ解析と全ゲノム情報復元
- 3) 全ゲノムリファレンスパネルの拡充
- 4) 全ゲノム解析の手法とその選択

### 4. オミックス解析等の実施計画

- 1) 多層オミックス参照パネルの構築と展開
- 2) 大規模メタボローム解析
- 3) エピゲノム解析
- 4) その他のオミックス解析等

### 5. 個別化予防に向けた統合解析のための基盤構築

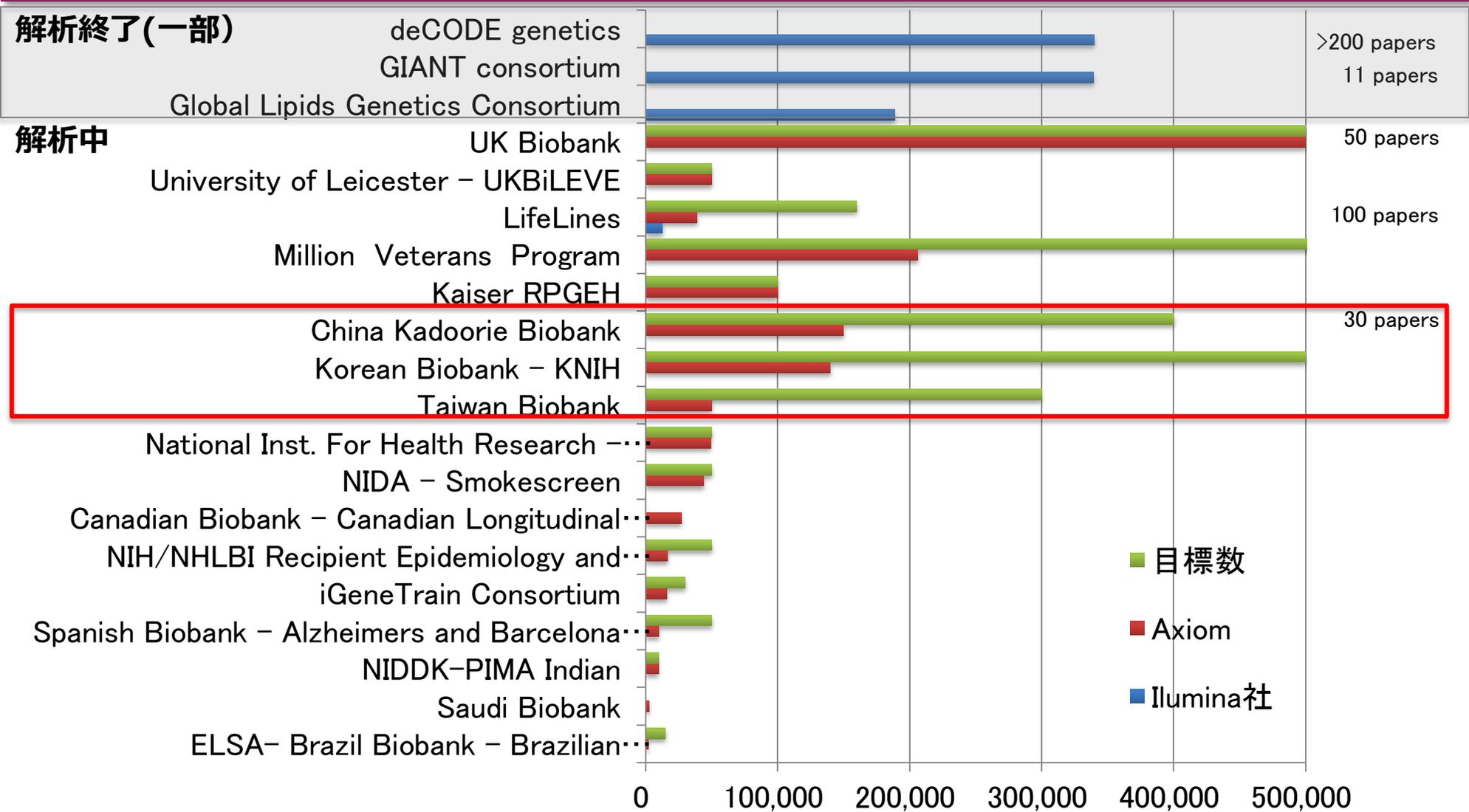
- 1) 発症リスク予測手法開発
- 2) 統合・知識データベースと病院連携

### 6. 遺伝情報回付

### 7. ゲノム医療実現のための人材育成

### 8. まとめ

# 海外のゲノムアレイ解析状況



欧米人ではすでに3,500万人以上のゲノムデータが蓄積している (Popejoy & Fullerton, Nature 2016)

**東アジア (中国、台湾、韓国)** は1-2年で数十万人規模に達する →**主導権確保が重要**

# 海外のゲノムアレイ・インピュテーション解析の現状

| バイオバンク                 | アレイ                               | インピュテーションリファレンス パネル                                     |
|------------------------|-----------------------------------|---|
| UK Biobank             | UK BiLEVE and UKB Axiom           | 1000 Genomes Project Phase 3 and UK10K panels (IMPUTE2) |
| LifeLines              | Axiom-NL                          | GoNL 1000 Genomes                                       |
| Kaiser RPGEH / UCSF    | Axiom (4 Different Ethnic Arrays) | 1000 Genomes Project Phase 1                            |
| Korea National Biobank | Axiom, Korean                     | Korean 1000 Genomes (probably)                          |
| Taiwan Biobank         | Axiom, Taiwanese                  | Taiwanese 1000 Genomes                                  |
| TMM Biobank            | Axiom, Japonica Array             | Japanese >3,400 Genomes                                 |

UK Biobankのインピュテーション解析では10万人規模の解析の段階で大きな成果を創出

喫煙と肺機能、COPD (50,008人) 15万人のデータより選択 喫煙歴有/無, FEV1により層別化 Lancet Resp Med 2015

認知機能と教育 (112,151人) Mol Psychiatr 2016

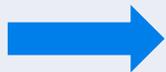
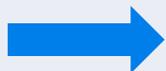
睡眠障害と代謝疾患、精神疾患の関連 (112,586人) Nat Genet 2017 (ahead of print)

他のコホート・バイオバンクデータと合わせた大規模解析や確認用解析の報告も多数

10万人規模の高精度推定全ゲノムデータが取得できれば、環境要因と遺伝要因の相互作用による疾患発症について、日本人の独自性を考慮した研究を推進できる

# ジャポニカアレイv2は疾患アレイの機能を強化

コストの大幅な削減と機能関連変異の追加

| Category                | Japonica v1 | Japonica v2   |
|-------------------------|-------------|---|
| タグSNP (X染色体含む)          | 638,269     | 638,269   |
| 薬剤応答関係 (ADME)           | 2,028       | 2,037   |
| Y染色体                    | 275         | 616   |
| ミトコンドリア                 | 70          | 104   |
| 過去にGWASで報告があったSNP       | 16,376      |  17,054  |
| HLA (免疫疾患などに関連)         | 3,906       |  6,914 |
| 機能を持つがImputationが難しいSNP | 3,990       | 3,990   |
| 合計                      | 659,253     | 659,328   |

# ジャポニカアレイv2 解析の意義

例) 高コレステロール血症に「関連する」病的変異

→ 日本人によるエビデンス不足により、DBによってアノテーションが異なる

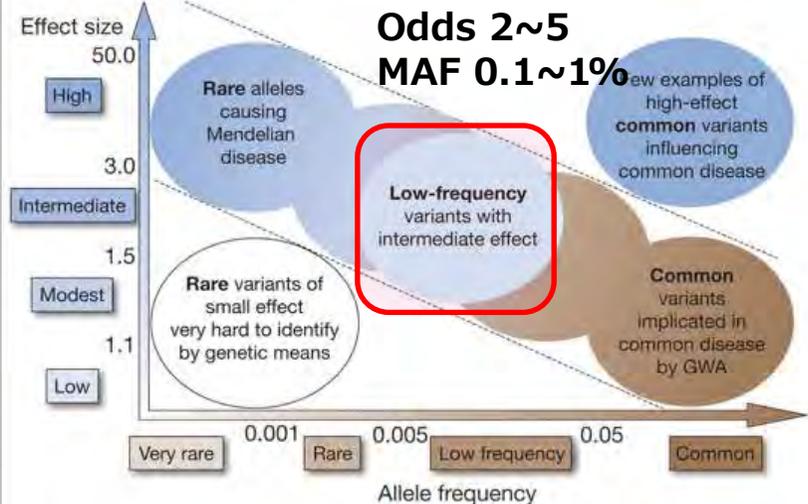
| ClinVar                | HGMD | SNP_ID                                | AA change | 2KJPN% | ExAC% | Publication  |
|------------------------|------|---------------------------------------|-----------|--------|-------|--|
| Likely pathogenic      | DM   | rs201102461<br>On JPA<br>(ジャポニカアレイ搭載) | R115H     | 0.39   | 0.02  | Chang 2003: 1/172人(AF0.0029) Taiwan<br>Chiou 2010: 1/102人(AF0.0049) Taiwan<br>Kim 2004: 1/31人(AF0.016) Korean<br>Yu 2002: 3/200人(AF0.0075) Jpn<br>Khoo 2000: 2/86人(AF0.012) Malaysia<br>Miyake 2009: 2/205人 (AF0.0049) Jpn<br>Mabuchi 2014: 5/489人(AF0.0051) Jpn<br>Chiou 2011: 1/125人(AF0.004) Taiwan   |
| Likely benign          | DM   | rs200990725<br>On JPA                 | R257W     | 0.05   | 0.007 | Nauck 2001: 1/56人(AF0.0089) German<br>Salazar 2002: 1/35人(AF0.014) Brazil (European)<br>Charng 2006: 1/51人(AF0.0098) Taiwan<br>Chiou 2010: 2/102人(AF0.0098) Taiwan<br>Shin 2015: 1/97人(AF0.0052) Korea<br>Fouchier 2005: 1/1177人(AF0.00042) Dutch<br>Santos 2014: 1/156人(AF0.0032) Brazil<br>Han 2015: 1hom/69人(AF0.014) Korea<br>Chiou 2011: 3/125人(AF0.012) Taiwan |
| Likely pathogenic      | DM   | rs746959386                           | L568V     | 0.02   | 0.002 | Hattori 1999: 1/120人(AF0.0042) Jpn<br>Shin 2015: 1/97人(AF0.0052) Korea<br>Miyake 2009: (7het+1hom)/205人 (AF0.022) Jpn<br>Mabuchi 2014: 4/489人(AF0.0041) Jpn<br>Han 2015: 1/69人(AF0.0072) Korea   |
| Uncertain significance | DM   | rs201971888<br>On JPA                 | D589N     | 0.05   | 0.006 | Fouchier 2005: 1/1177人(AF0.00042) Dutch<br>Charng 2006: 1/51人(AF0.0098) Taiwan<br>Chiou 2010: (1het+1hom)/102人(AF0.015) Taiwan<br>Shin 2015: 1?/97人(AF0.0052) Korea<br>Han 2015: 1hom/69人(AF0.014) Korea<br>Chiou 2011: 3/125人(AF0.012) Taiwan   |
| Likely pathogenic      | DM   | rs201573863<br>On JPA                 | P699L     | 0.02   | 0.002 | Schuster 1995: 1/10人(AF0.05) German<br>Thiart 2000: 1/16人(African) (AF0.031)<br>Van Gaal 2001: 1/98人(AF0.0051) Belgian<br>Fouchier 2001: 1/1641人(AF0.00030) Dutch<br>Bertolini 2013: : 1/1048人(AF0.00048) Belgian<br>Sharifi 2016: 2/161人(AF0.0062) Italy  |
| NA                     | DM   | rs13306505                            | A860V     | 1.00   | 0.007 | Miyake 2009: 2/205人 (AF0.0049) Jpn<br>Do 2015: 1/4705人若年MI (AF0.00011) US  |

# インピュテーションを活用した場合の検出力推定

Rare variants検出にはコントロール群のサイズ向上が有効

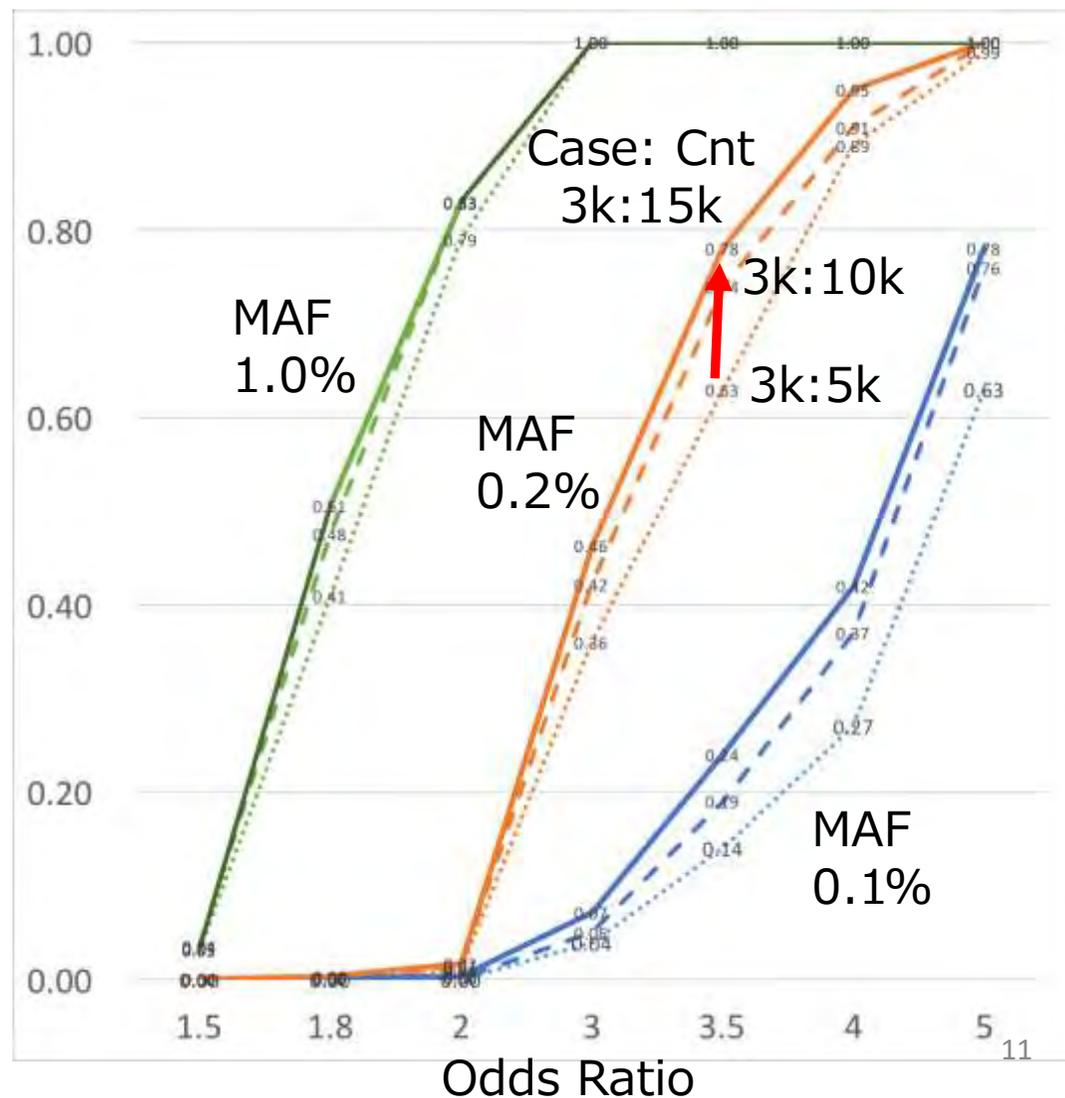
効果のある低頻度変異が多因子疾患のターゲット変異

Odds 2~5  
MAF 0.1~1%



有病率を考慮し、12~15万人のゲノムデータを準備し  
疾患に応じて質の高い10万人規模コントロール群を提供

差のある変異の推定検出割合

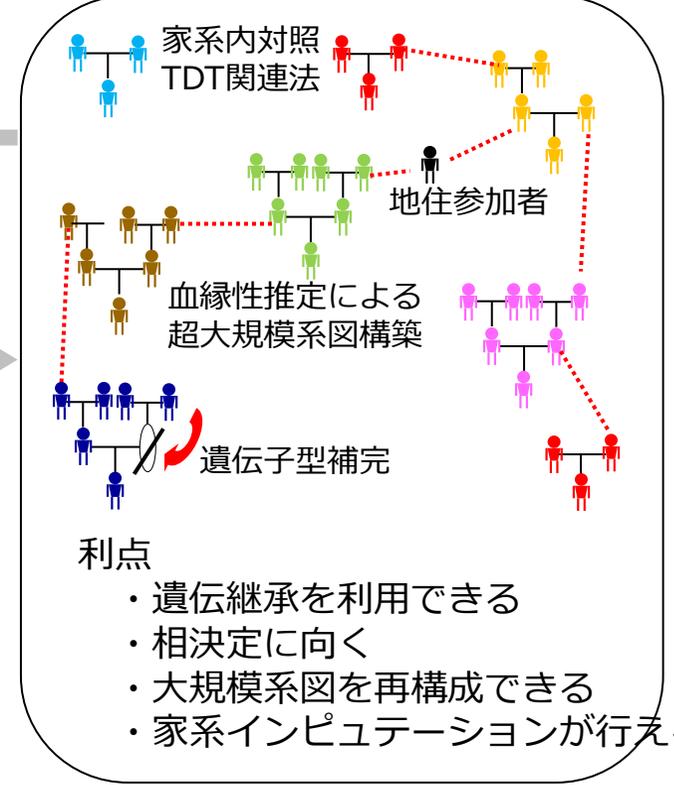


# 遺伝子環境相互作用への挑戦

## 地域住民コホート



## 三世代家系コホート



相乗的効果

環境要因

遺伝子と環境  
相互作用

遺伝要因

対象疾患の  
要因分布

循環器疾患

アレルギー性疾患・うつ病

自閉症

地域住民コホートの得意とする範囲

未知の領域  
(相互作用)

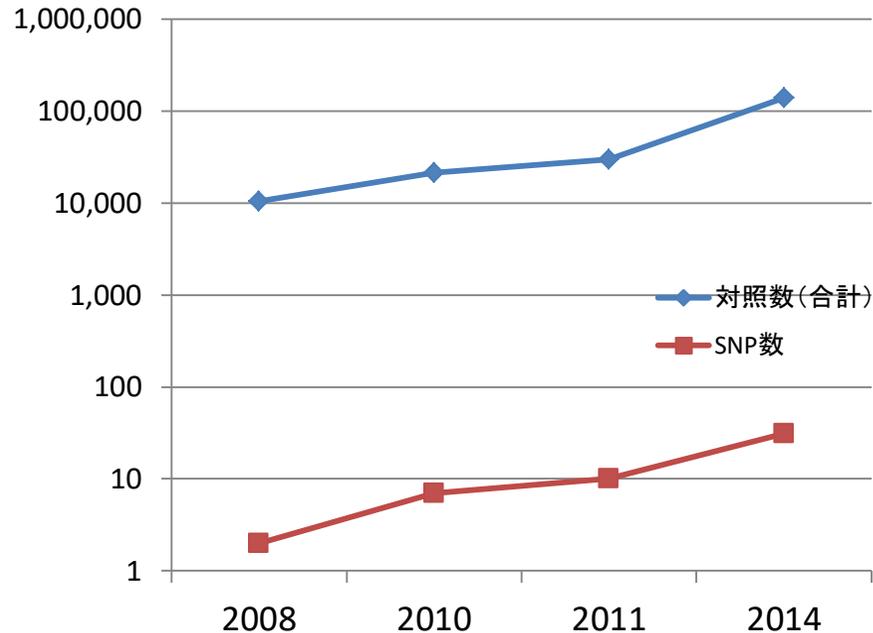
三世代家系コホートの得意とする範囲

# コントロールの必要数

|                      | 疾患               | 症例数     | 対照数     | コホート/コンソーシアム  | 発表年                        |
|----------------------|------------------|---------|---------|---|----------------------------|
| 肝疾患                  | アルコール性<br>肝硬変    | 1,800   | 2,300   |   | Nat Genet,<br>2015         |
| がん                   | 前立腺がん            | 25,000  | 24,000  | The Institute of Cancer Research など                                 | Nat Genet,<br>2013         |
|                      | 子宮内膜癌            | 7,700   | 37,000  | Australian National Endometrial<br>Cancer Study (ANECS) など          | Nat Genet,<br>2016         |
| 呼吸器疾患                | 慢性閉塞性肺<br>疾患     | 50,000  |         | UK BioBank, UK BiLEVE   | Lancet Respir<br>Med, 2015 |
| 努力性肺活量               | 量的形質             | 85,000  |         | 26 施設 (ARIC, Lifelines, など)   | Nat Genet,<br>2014         |
| 認知機能と体力、メンタルヘルス      | 量的形質             | 112,000 |         | UK BioBank<br>24 GWAS consortia                                     | Mol Psychiatry,<br>2016    |
| 心血管障害                | 心筋梗塞、冠<br>動脈狭窄など | 61,000  | 124,000 |   | Nat Genet,<br>2015         |
| 脳血管障害                | 脳梗塞              | 24,000  | 132,000 | Neurology Working Group of<br>CHARGE Consortium                     | Lancet Neurol,<br>2016     |
| 糖尿病                  | 2型糖尿病            | 27,000  | 135,000 |   | Nature, 2016               |
| 出生時体重と糖尿病、<br>心血管生障害 |                  | 150,000 |         | UK BioBank  | Nature, 2016               |
| コレステロール値、<br>中性脂肪値   | 量的形質             | 189,000 |         | Global Lipids Genetics Consortium                                   | Nat Genet,<br>2013         |
| 収縮期・拡張期血圧、<br>脈圧など   | 量的形質             | 320,000 |         | GERA cohort, Consortium for Blood<br>Pressure (ICBP) and UK Biobank | Nat Genet,<br>2016         |

# GWAS解析における検体数の影響

## 2型糖尿病

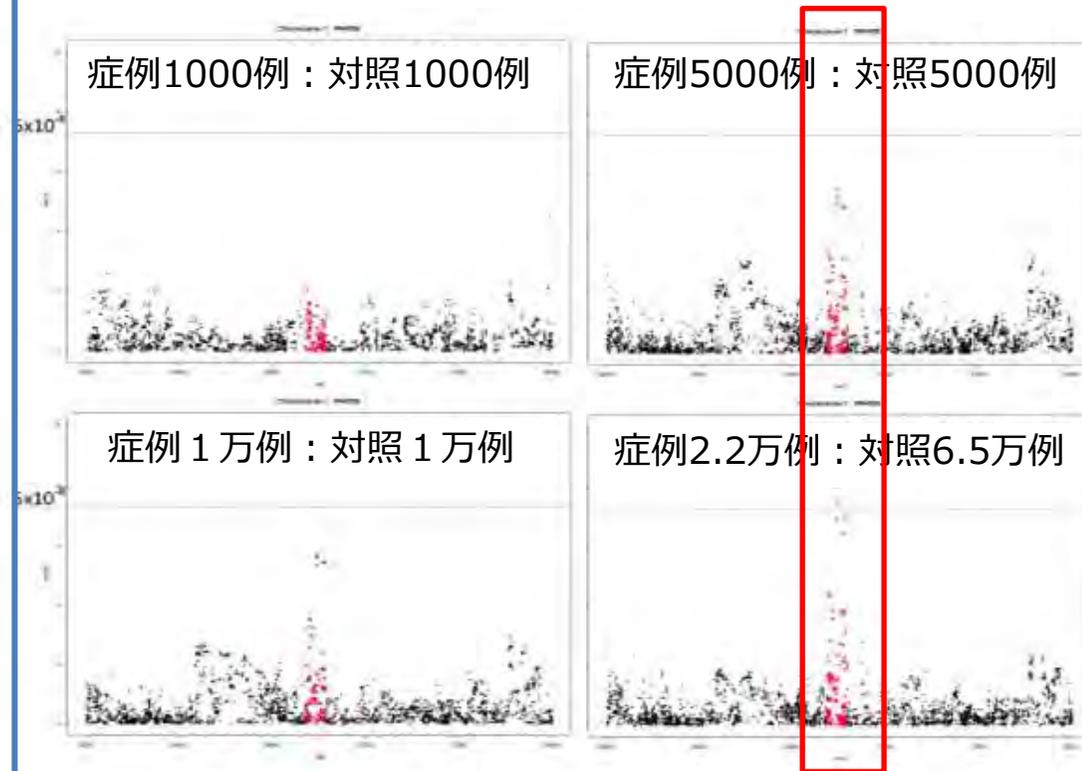


検体数増加とともに有意なSNP数は増える

| 発表年        | 症例数(合計) | 対照数(合計) | SNP数 |
|------------|---------|---------|------|
| 2008*      | 9,569   | 10,361  | 2    |
| 2010*      | 22,142  | 21,283  | 7    |
| 2011*      | 25,079  | 29,611  | 10   |
| 2014*      | 47,979  | 139,611 | 31   |
| SNP数との相関係数 | 0.983   | 0.991   | 1    |

理研・東大の共同研究の時系列データ \*Nat Genet.

## 心血管障害



Cardiogram consortium, Schunkert et al. 2011; Nat Genet.

総計**8.7万**でようやく有意差が出る

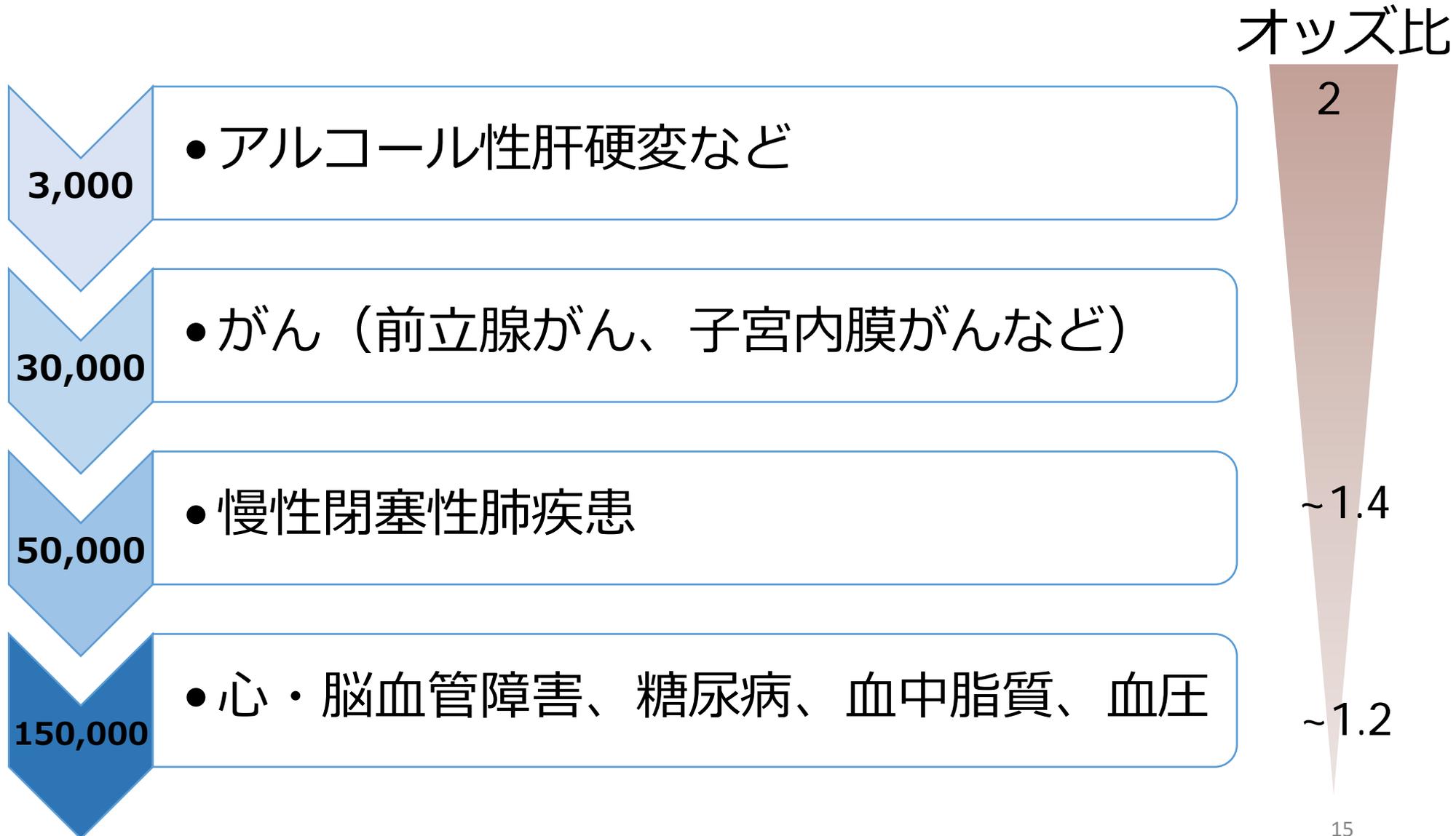
横軸：染色体1p32.2の領域

縦軸：疾病とSNPの関連の強さ

グラフの横線：統計的有意の閾値

赤い点はこの研究で関連するとされた遺伝子上のSNP

# 疾患解析で要求されるコントロール数



# ゲノム・オミックス解析

## 第2段階計画概要

---

### 1. はじめに

### 2. 大規模ゲノム解析の意義

- 1) 海外のコホート・バイオバンクにおけるゲノム医学研究の現状、大規模アレイ解析との比較
- 2) 疾患関連遺伝子研究に向けた基盤の提供
- 3) 大規模ケース・コントロールスタディでの活用

### 3. ゲノム解析の実施計画

- 1) ゲノム解析のロードマップ
- 2) 大規模SNPアレイ解析と全ゲノム情報復元
- 3) 全ゲノムリファレンスパネルの拡充
- 4) 全ゲノム解析の手法とその選択

### 4. オミックス解析等の実施計画

- 1) 多層オミックス参照パネルの構築と展開
- 2) 大規模メタボローム解析
- 3) エピゲノム解析
- 4) その他のオミックス解析等

### 5. 個別化予防に向けた統合解析のための基盤構築

- 1) 発症リスク予測手法開発
- 2) 統合・知識データベースと病院連携

### 6. 遺伝情報回付

### 7. ゲノム医療実現のための人材育成

### 8. まとめ

# 全国のゲノム医療研究のゲノム基盤構築にむけたロードマップ ～全ゲノムリファレンスパネルの拡充とデータシェアリング～

- ◎ 8,000人に向けた人数規模の拡充による新規レアバリアントの更なる収集と全ゲノムリファレンスパネルの構築と公開による国内外の遺伝性疾患の原因同定の加速を実施
- ◎ 参照とする国際参照配列および日本人の配列を考慮した全ゲノムリファレンスパネル定期的な改定による高精度化を実施
- ◎ 全ゲノムリファレンスパネルの質的向上のための疾患関連難読化領域のハプロタイプの確定や三世代を最大限活用した解析を実施

全国の研究者の利用(国外も含む)による遺伝性疾患(単一、多因子ともに)の理解の加速  
(例)小児IRUD 成人IRUD 遺伝性がん 国内外ゲノムコホート事業など

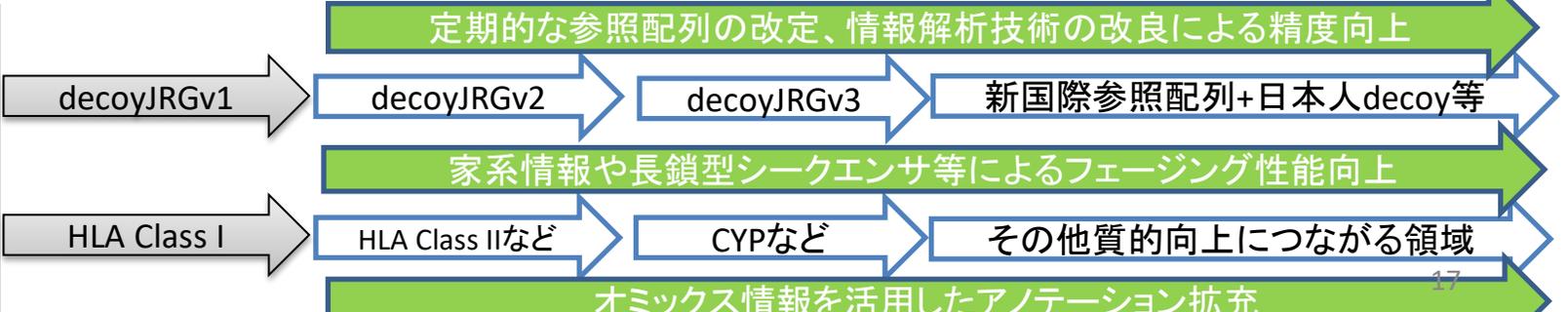
## 全ゲノムリファレンスパネル情報の提供

| 全ゲノム解析                         | H28年度  | H29年度  | H30年度  | H31年度  | H32年度  |
|--------------------------------|--------|--------|--------|--------|--------|
| シーケンス人数                        | 4000人  | 5000人  | 6000人  | 7000人  | 8000人  |
| 全ゲノムリファレンスパネルの人数拡大によるレアバリアント拡充 | 3800万個 | 4400万個 | 4950万個 | 5500万個 | 6000万個 |
| レアバリアント総数※予測                   |        |        |        |        |        |

アレイの全ゲノム復元性能の向上



全ゲノムリファレンスパネルの質的向上



decoyJRGv1

decoyJRGv2

decoyJRGv3

新国際参照配列+日本人decoy等

HLA Class I

HLA Class IIなど

CYPなど

その他質的向上につながる領域

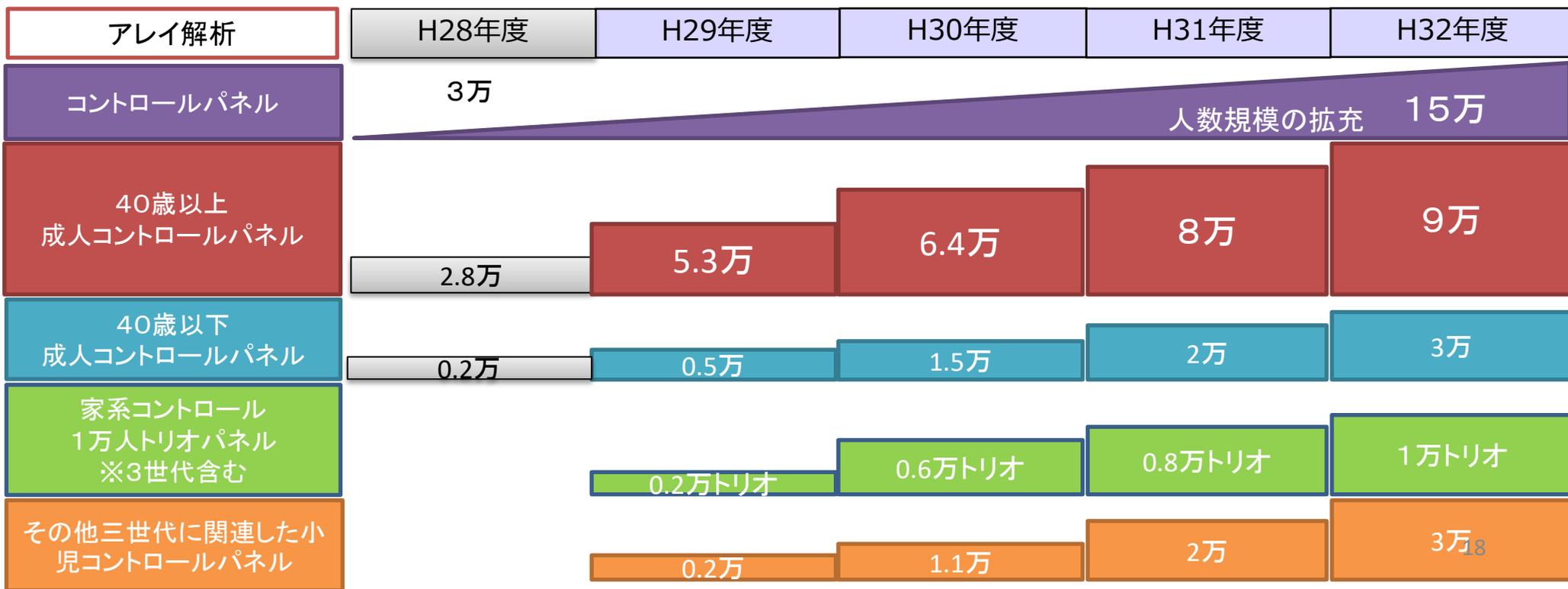
# 全国のゲノム医療研究のゲノム基盤構築にむけたロードマップ ～SNPアレイコントロールパネルの拡充とデータシェアリング～

## ●第2期(平成29年度～平成32年度)の重点目標と考え方

- ◎ コントロールパネルとして①40歳以上の**成人パネル(主に60歳以上)**②40歳以下の**成人パネル**③**1万トリオパネル**等を構築
- ◎ 2型糖尿病、高血圧などある程度高年齢で発症する多因子疾患については年齢層が高いコントロールが有効であることから**①のコントロールパネルを優先して構築**
- ◎ 機構独自の三世代コホート調査における**家系情報を活用し疾患と少なくとも同数の家系情報をコントロールとして構築**をする。またコントロールの構築においては、トリオが形成されている家系を優先

大規模健常コントロールとしての活用による疾患コホート研究における多因子疾患を中心としたゲノム医療研究の理解の加速 (例)2型糖尿病研究、高血圧、B型肝炎など

### 全ゲノムリファレンスパネルを用いた15万人の全ゲノム復元と提供



# TMM計画重点疾患のゲノム基盤構築に向けたロードマップ

## TMM計画重点疾患のゲノム基盤情報構築と全国の研究者へのデータシェアリング ～重点疾患のゲノムリスク要因の同定と疾患リスク予測精度の向上～

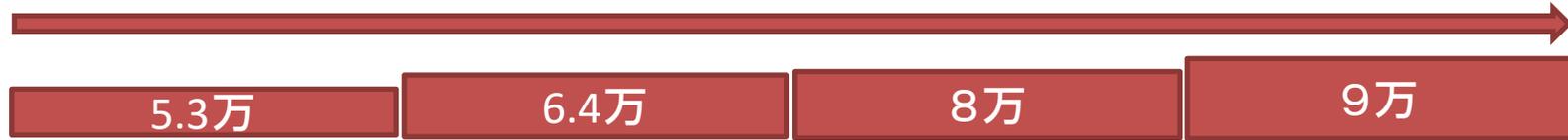


### SNPアレイ情報 ※家系情報含む

- ◎ 各年度に重点疾患を設定、設定する課題の順序は前向きコホート調査でアウトカムが取得ができ、成果を創出できる順に実施
- ◎ 取得順序をアウトカムの順序にあわせることで、コントロールパネルと疾患関連遺伝子解析のための基盤形成を同時に実現
- ◎ 各課題の疾患同定に有効な場合には、三世代コホート調査の特長を最大限活用し家系解析を実施
- ◎ 各課題対象者から全ゲノム収集した新規レアバリエントをSNPアレイで大規模に取得する連携体制を構築し原因バリエントを同定

#### ① 検査値関係

検査値異常(慢性閉塞性肺疾患、脳卒中・心筋梗塞・狭心症などの因子含む)



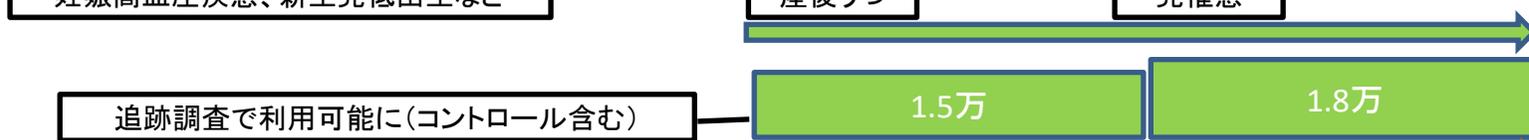
#### ② 妊婦疾患

妊娠高血圧症・低出生体重など



#### ③ 追加調査関係

脳形態変化、認知機能低下、要介護認定など



#### ④ 小児の疾患

ADHD、自閉スペクトラム、先天形態異常、先天性代謝異常、アトピー性皮膚炎など



### 全ゲノム情報 ※家系情報含む

- ◎ 疾患・形質の対象者を追加で全ゲノム解析し、全ゲノムリファレンスパネルには含まれない、新規疾患レアバリエントを効率良く収集

疾患関連レアバリエント拡充

100人～必要数

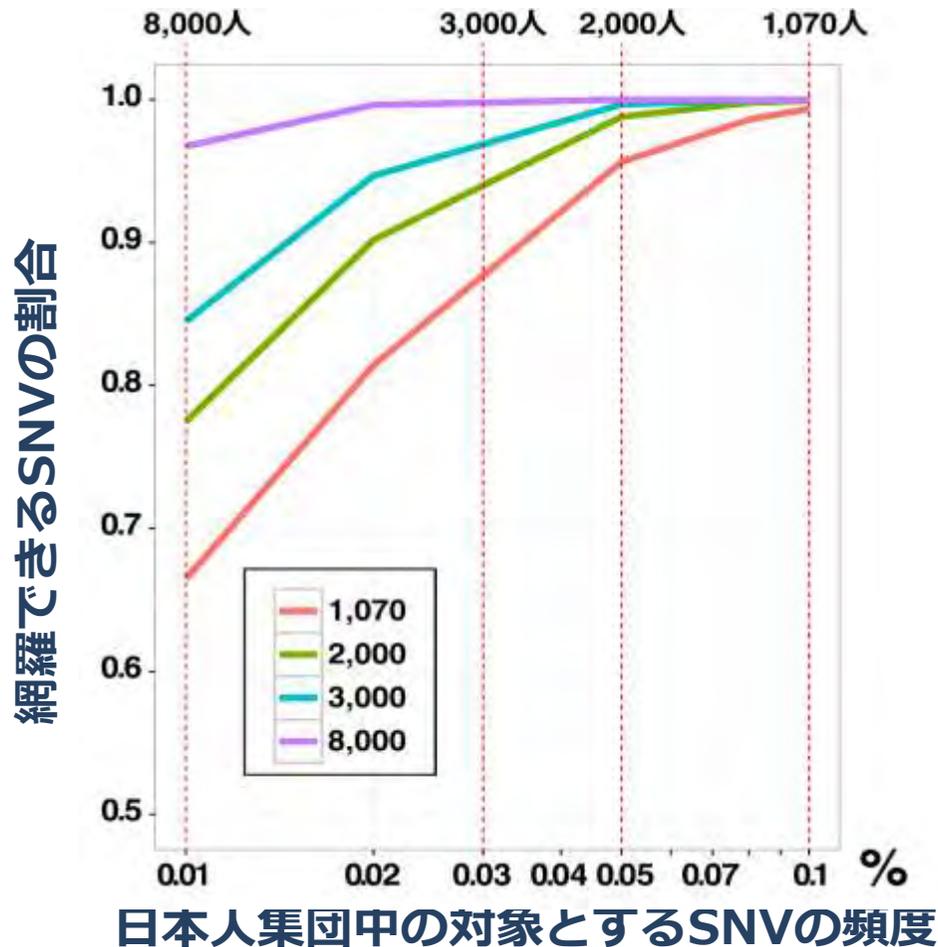
100人～必要数

100人～必要数

100人～必要数

# 日本人全ゲノムリファレンスパネルを 8,000人を目標とすることの意義

## 人数の拡充と網羅できるSNVの頻度



| 検体数   | 網羅できるSNVの頻度 |
|-------|-------------|
| 89    | >1%         |
| 1,070 | >0.1%       |
| 2,000 | >0.05%      |
| 3,000 | >0.03%      |
| 8,000 | >0.01%      |

**優性遺伝だと  
1万人に1人までの変異情報を  
ほぼ網羅**

先天性難聴 1万人出生に5人  
 血友病 1万人出生に1人  
 ベッカー型筋ジストロフィー 2-3万人に1人

# リファレンスパネルの高精度化計画案

高深度全ゲノム解析



【成果】被災東北地区の遺伝要因による疾患解明に向けた0.5%以上のSNVを網羅する全ゲノムリファレンスパネルの構築と公開

第一段階

|        |                   |           |
|--------|-------------------|-----------|
| H27.12 | 1,070人パネル公開       | 2,100万SNV |
| H28.6  | 2,049人パネル公開       | 2,800万SNV |
| H29.3  | 3,300人のパネル公開 (予定) |           |

【目標】

- ①日本人の全ゲノムリファレンスパネルの拡充と公開
- ②情報解析技術の開発による全ゲノムリファレンスパネルの質的向上
- ③不死化リンパ球、オミックス解析によるアノテーション拡充

第二段階

|       |                                     |
|-------|-------------------------------------|
| H30.3 | 年1000人の全ゲノム解析 (他地域含む)、機能解析アノテーション充実 |
| H33.3 | 8,000人以上のパネル機能解析アノテーション公開           |

【効果】日本人のもつレアなSNV情報取得と疾患原因変異の絞込みへの活用

疾患の方がもつ変異候補 300万

| リファレンスパネル | 変異候補     |
|-----------|----------|
| 1000人     | 数個～50個程度 |
| 2000人     | 数個～25個程度 |
| 8000人     | 1～数個程度   |

※ より稀なSNVに疾患に影響が高い変異がある  
稀なSNVは日本人集団に特異的なSNVが多い

【効果】 SNPアレイ(66万SNP)のみで個人毎の0.1%以上のSNP(約300万)の高精度復元

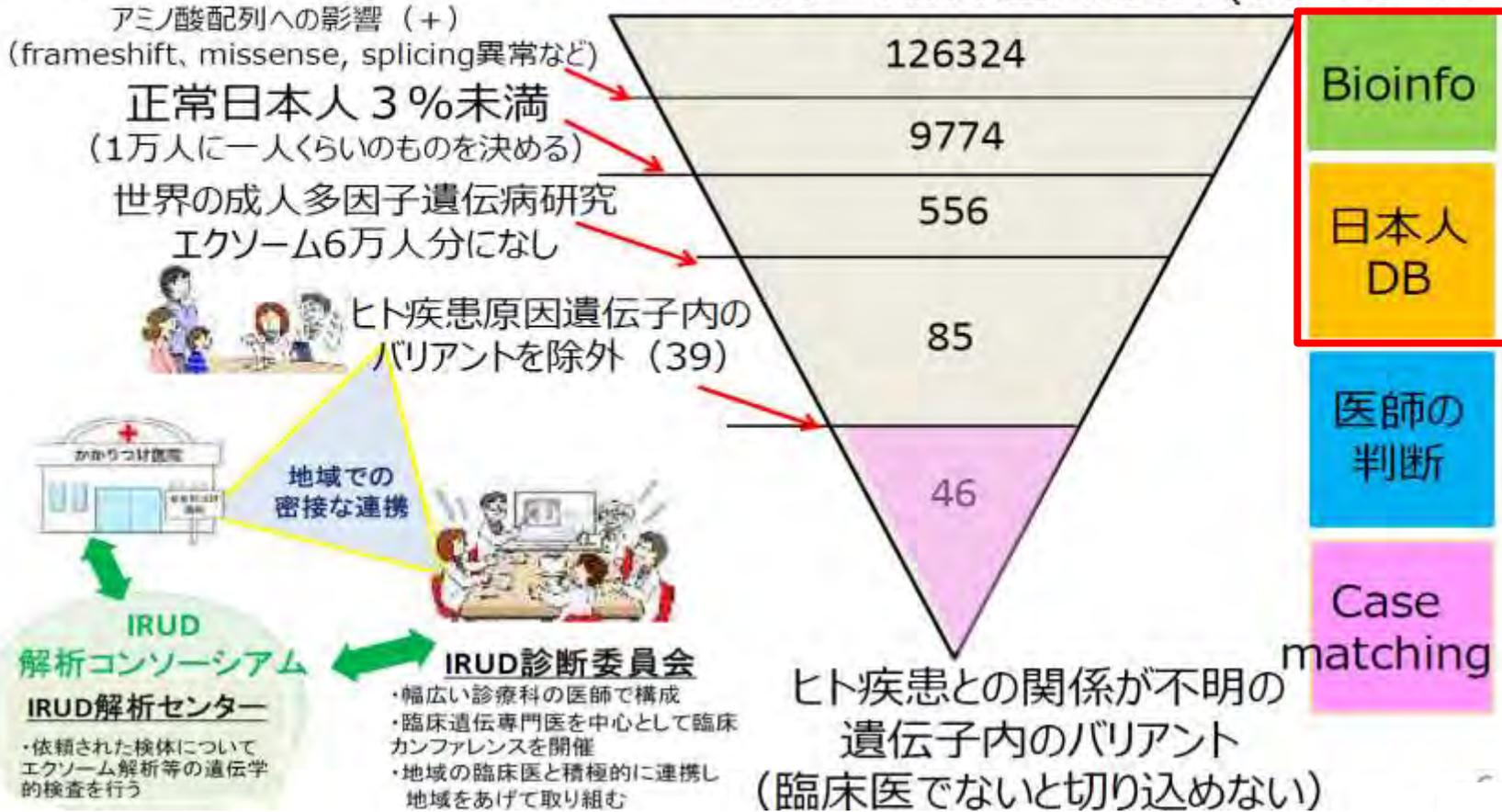
# データシェアが加速するゲノム医療 (IRUD)

## ToMMo2KJPNによる未診断疾患の原因変異絞り込み

### Phenotype is everything

エクソーム由来全バリエント(例)

必要な  
リソース

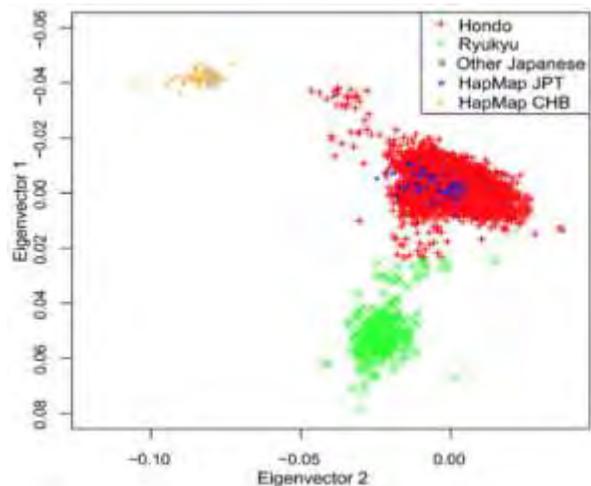


ToMMo  
2KJPN  
活用

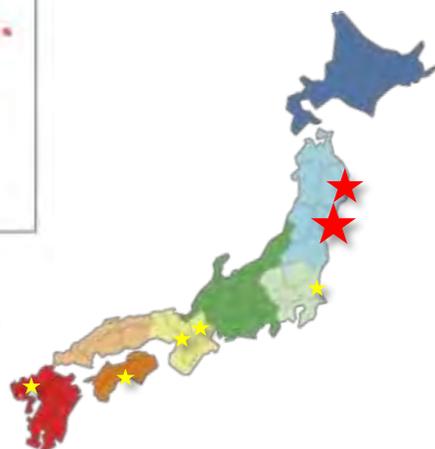
本年度からIRUD解析センターとしても参画

# 他地域検体の取り込みによる高精度化

## 日本人集団内での多様性



Yamaguchi et al, AJHG, 2008



コホート間連携による  
他地域サンプルの取得

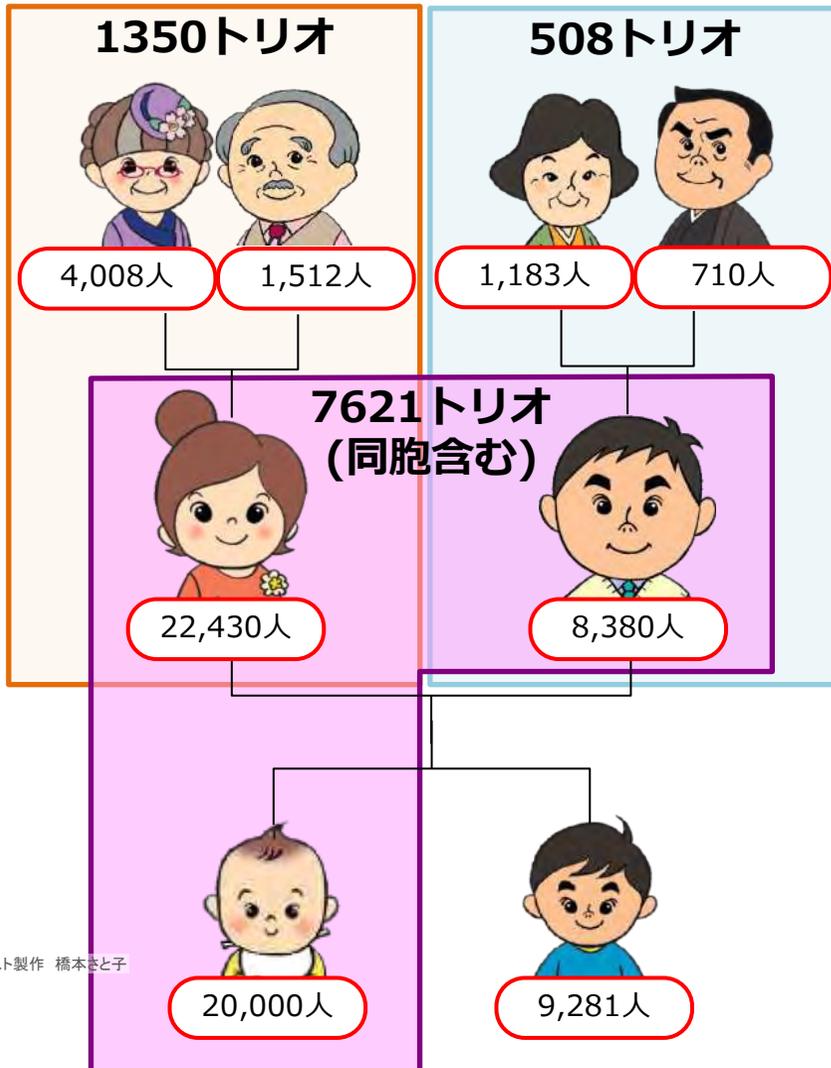
## 解析状況

| コホート        | 解析検体数 |
|-------------|-------|
| ToMMo他地域検体  | 約400  |
| ながはま (近畿)   | 約400  |
| 長崎医療セ(九州)   |       |
| JPHC (四国)   |       |
| J-MICC (中部) |       |
| 大阪大 (近畿)    |       |
| 総計          | 約800  |

### インピュテーション性能のためのパネル高精度化戦略

- 量的対策：全ゲノム数の増強（8000人を目指す）
- 質的対策：他地域サンプルを入れた多様性の向上

# TTT (10k Trios) の解析 三世代コホート



## <TTT; 10k Trios>

- 約7万人の三世代コホートから約1万組の親子トリオをピーリングできる

## <TTT; 10k Trio利用の意義>

- トリオはde novo mutationの直接観察を可能にする

(Conrad et al. Nat Genet 2014等)

- 特に、三世代への遡上によって父系と母系の突然変異を区別できる

(Kong et al. Nature 2012等)

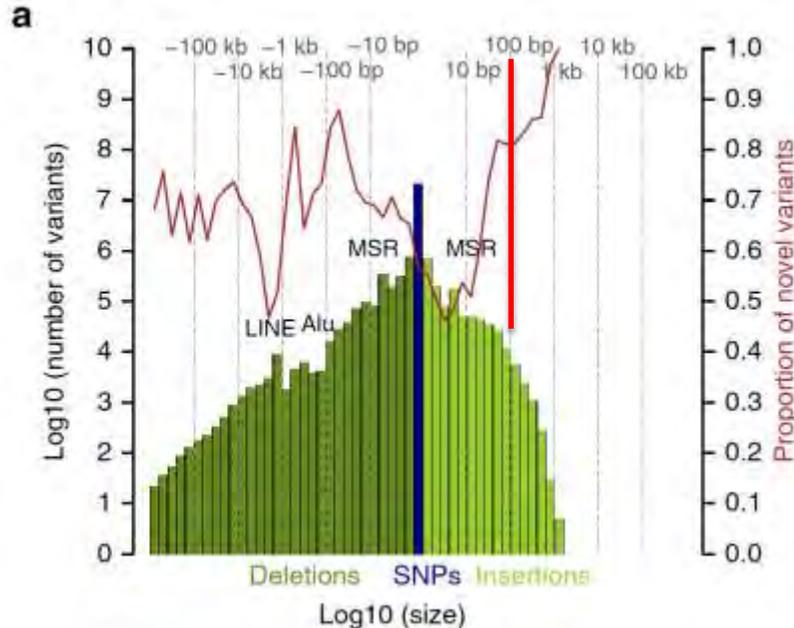
- Imputation用ハプロタイプコレクションを高精度化できる

(HapMap3, Nature 2010; 1000GPC, Nature 2012等)

- TDT等の家系内対照デザインを可能にする
- 上記のようなゲノム医学への種々の強力なツールを提供するリソースとなる

# 挿入変異の検出が今後の課題

従来法は100bp以上の挿入欠失やSTRを十分に捕捉できない



SNPと同様に挿入欠失やSTRも  
生物学的に大きな影響をもたらす

例 神経変性疾患 ポリグルタミン病(CAG)<sub>n</sub>  
脆弱X症候群 (CGC)<sub>n</sub>  
アンドロゲン受容体と発癌 (CAG)<sub>n</sub>

正常 ~30リピート (~90bp)  
疾患 50~100リピート (150 ~300bp)



長鎖リードデータが必要だが、PacBioは多検体の解析には適さないが  
259PE新規手法による解析は選択肢となりうる

民間では実施していない解析プロトコルで精度の高いシーケンス情報を取得し、  
包括的なゲノム解析、フェージング性能を向上させるプロトコルの開発を実施する

# ゲノム・オミックス解析

## 第2段階計画概要

---

### 1. はじめに

### 2. 大規模ゲノム解析の意義

- 1) 海外のコホート・バイオバンクにおけるゲノム医科学研究の現状、大規模アレイ解析との比較
- 2) 疾患関連遺伝子研究に向けた基盤の提供
- 3) 大規模ケース・コントロールスタディでの活用

### 3. ゲノム解析の実施計画

- 1) ゲノム解析のロードマップ
- 2) 大規模SNPアレイ解析と全ゲノム情報復元
- 3) 全ゲノムリファレンスパネルの拡充
- 4) 全ゲノム解析の手法とその選択

### 4. オミックス解析等の実施計画

- 1) 多層オミックス参照パネルの構築と展開
- 2) 大規模メタボローム解析
- 3) エピゲノム解析
- 4) その他のオミックス解析等

### 5. 個別化予防に向けた統合解析のための基盤構築

- 1) 発症リスク予測手法開発
- 2) 統合・知識データベースと病院連携

### 6. 遺伝情報回付

### 7. ゲノム医療実現のための人材育成

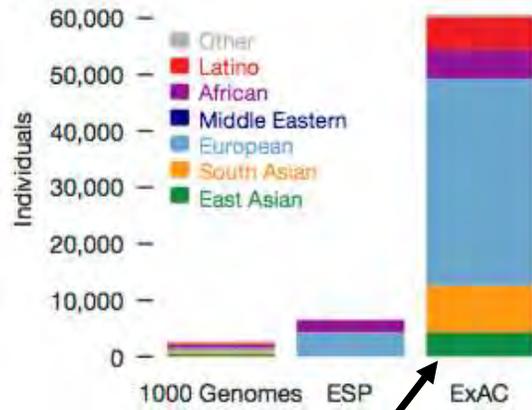
### 8. まとめ

# 細胞解析基盤構築とマルチオミックス解析

Nature 536, 285–291 (18 August 2016)

## ARTICLE

### Analysis of protein-coding genetic variation in 60,706 humans



東アジア系の変異情報の不足  
→ ToMMoによる変異情報の蓄積

NHGRI GWAS catalog 中の日本人データ

|     | SNP数       | 形質・疾患数     |
|-----|------------|------------|
| 日本人 | 350 (4.4%) | 60 (10.4%) |
| 全体  | 7916       | 574        |

東アジア系の変異アノテーションも不足  
→ All Japanでの機能解析体制の構築

## TMMの機能解析への取り組み

### 細胞解析基盤形成による機能解析

バイオバンク保存血液試料  
Bリンパ球      Tリンパ球



EBV感染



不死化B細胞

(現状800人分)

IL2+抗CD3

増殖T細胞

(現状1070人分)

ゲノム情報が明らかな数千人分の  
各種細胞試料の蓄積

試料分譲

細胞試料・解析情報の分譲によるゲノム医療基盤強化

情報分譲

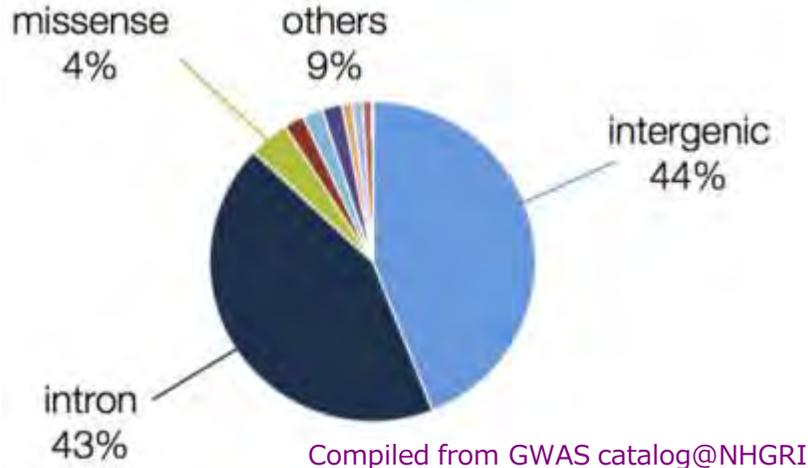
薬剤に対する反応性、環境因子への応答  
(酸化ストレス負荷など)、  
エピゲノム・トランスクリプトーム・メ  
タボローム解析によるデータ取得

TMMでの解析基盤を利用した解析

# 多型の機能解析における発現情報の重要性

GWASにより関連が報告されている変異の大部分はintergenic/intronにある

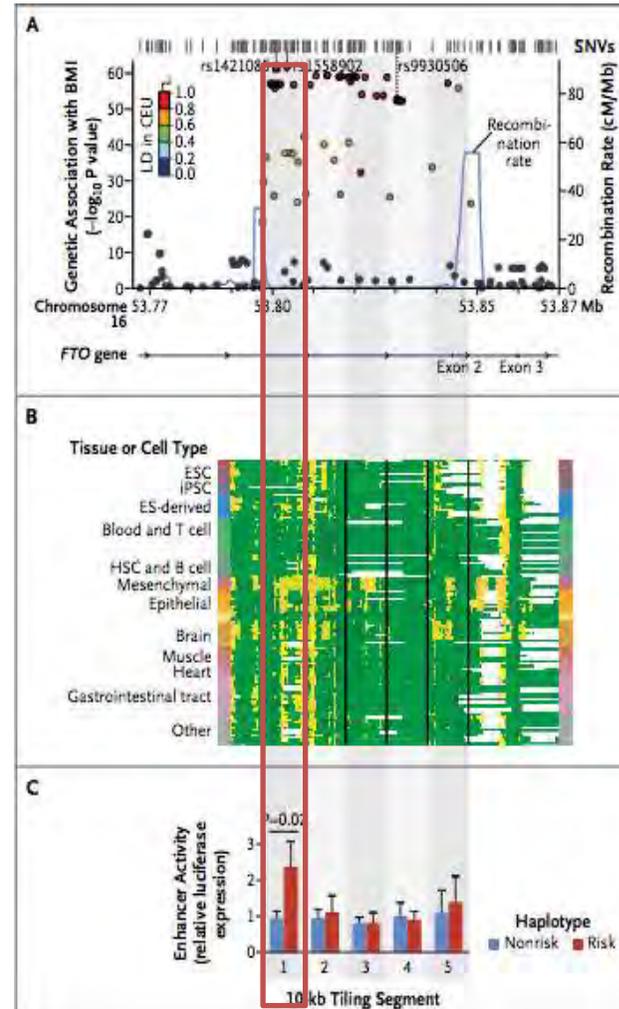
- 遺伝子機能での解釈の難しさ
- **細胞解析基盤を活用し遺伝子発現への影響を評価し、変異の絞り込み**



## 機能解析基盤でのRNA-seqの利点

- 発現している遺伝子の構造がわかる
- 変異による遺伝子発現の変化がわかる
- 遺伝子共発現から機能関連遺伝子もわかる
- 変異の絞り込み
- **日本人での変異アノテーション充実**

## 発現量情報とGWASを合わせて変異絞り込み



# GWAS

変異の絞り込みが出来てない

**エピゲノム**  
細胞・組織でのクロマチンの状態

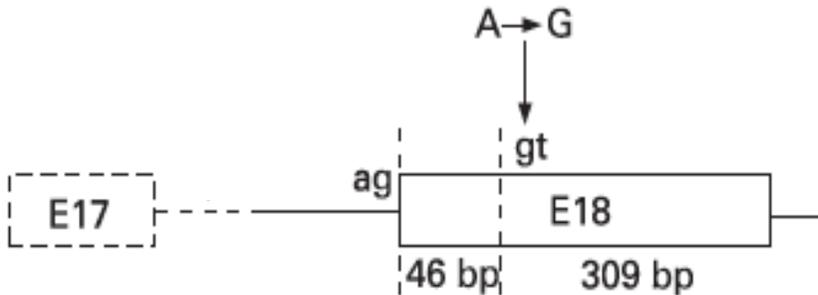
**遺伝子発現**

Claussnitzer et al, NEJM, 2015

# 発現情報に基づく多型の機能解析を個別化医療へつなげる

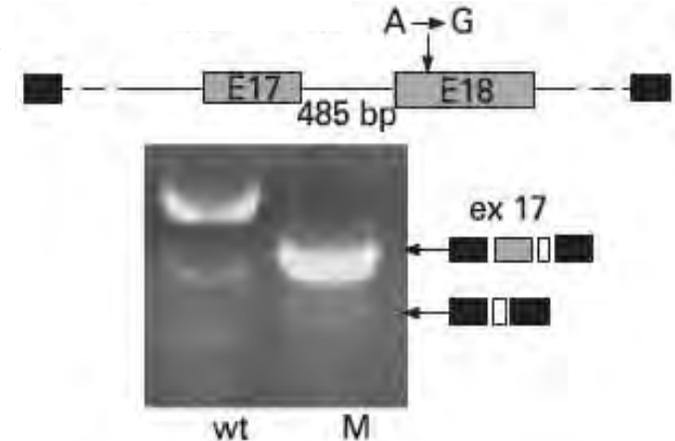
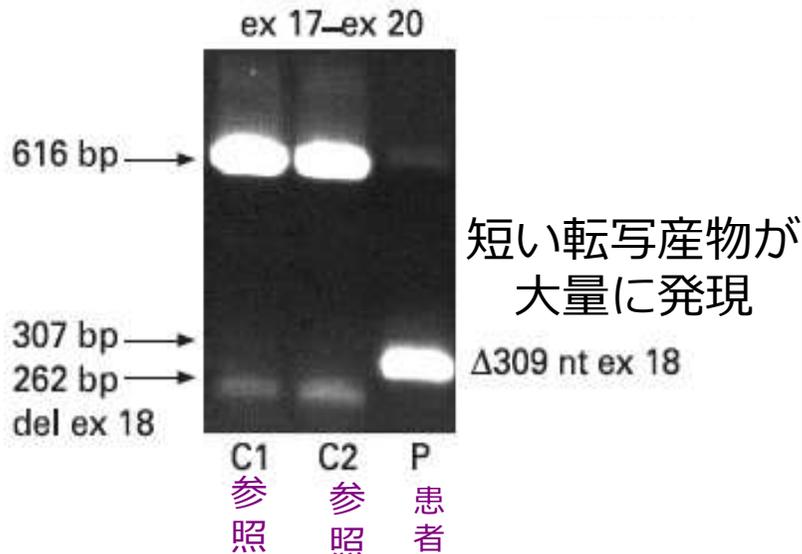
乳がん関連遺伝子 BRCA2: rs32937362 (A>G:I2675V)

情報解析では、Ile→Valという重篤には思えない変異に見える



培養細胞での組み替えDNAを用いた検証により変異がプライス異常を引き起こし、機能が失われていたことが判明

患者 (P) のリンパ球でのRT-PCR



情報解析のみでこの種の機能に影響を持つ変異を検出することは困難でありRNA-seqによる遺伝子構造のカタログ化が必須

## 第二段階のメタボローム解析の目標

- TMM前向きコホート研究は追跡調査をもとに、疾患の原因となる遺伝要因と環境要因の相互作用を明らかにすることを目標とする。
- 個人の代謝環境（メタボローム）はセントラルドグマの最下流にあり、遺伝要因と環境要因のエンドポイントである。このため、様々な遺伝要因・環境要因と疾患発症などの表現型との因果関係を解明する上で最も効果的な解析対象である。
- しかし、関係する要因がゲノム以上に膨大で、かつ複雑に絡み合うため、個々の要因を遡上して明らかにするためには、従来のゲノムを主とした症例対照研究と同等以上のサンプル数の解析が必要となる。
- このためToMMoの解析基盤の中で特に定量性が高いNMR法を最大限に活用して大規模メタボローム解析を行う。さらに、高感度な質量分析法を対象を絞って活用することでより詳細な解析を進めるとともに国際標準の確立に貢献する。

### 第二段階の課題

1. 大規模メタボロームGWAS解析を行い、各ゲノム多型がタンパク質の機能や量に与える影響（最終的に個人の体質に反映）を代謝物の量的変化から明らかにする。
2. 追跡調査において早期に発症が確認できる疾患を対象として症例対照研究を行い、疾患関連マーカーを探索することで、個別化医療・予防に貢献する。
3. 1と2の代謝物情報と調査表情報等から、生活習慣等の環境要因が個人の体質に与える影響を解明する。

以上の3課題の成果を元に高精度版参照パネルを構築しデータシェアリングを行う。

## 第二段階の解析計画

### 課題1 大規模メタボロームGWAS解析による遺伝要因の解明

詳細検査情報のあるコホート参加者（最大5万人）を対象に大規模メタボローム解析を行い、得られた代謝物情報とゲノム情報との関連解析（GWAS）を行うことで、各ゲノム多型の機能解明をおこなう。特に全ゲノムインピュテーションによる、低頻度バリエーションも対象とする大規模な解析を行うことでその効果を明らかにする。

### 課題2 早期発症疾患のメタボローム解析

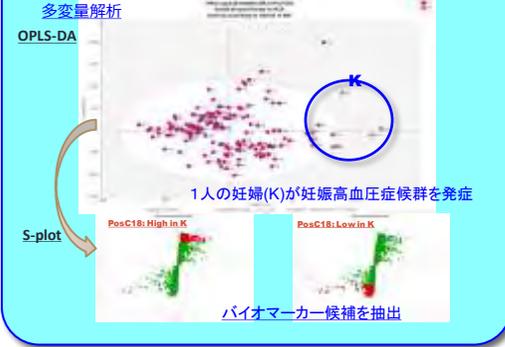
疾患症例の中でも発症が早期に判明する三世代コホートの症例（妊娠高血圧症候群等。発症率合計約20%）を中心にメタボローム解析を行い、病態の解明と早期診断マーカーの探索を行なう。さらに追跡調査で発症が確認された疾患も対象とする。合計で2万人を想定する。

### 課題3 環境要因が疾患発症に及ぼす影響の解明

課題1と2のメタボローム解析の成果（合計最大7万人）と、各種コホート情報（調査票、詳細生理学的検査、MRI等）との関連解析により、個人の代謝環境に影響を与える環境要因を明らかにし、遺伝要因と合わせて、疾患（生活習慣病等）発症等の表現型に影響を与える因子の全容を明らかにする。

# 第二段階 大規模メタボローム解析の解析計画

## 研究課題1 妊娠高血圧症候群のマーカー探索



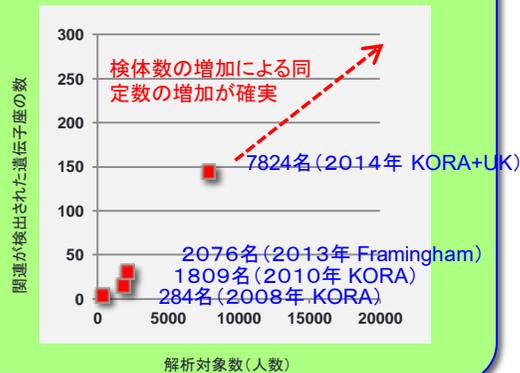
疾患発症が早期に判明する三世代コホートの症例（妊娠高血圧症候群等。発症率合計約20%）を中心にメタボローム解析を行い、病態の解明と早期診断マーカーの探索を行なう。

## 研究課題2

症例対照研究  
(最大2万人)

個別化医療の推進

## メタボロームGWAS解析における同定遺伝子座の推移



## 研究課題3

環境要因の影響の解析  
(合計 最大7万人)

表現形の確立  
個別化予防の実現

課題1, 2のメタボローム解析の成果と、各種コホート情報（調査票、生化学検査、MRI等）との関連解析により、個人の体質に影響を与える環境要因を明らかにし、遺伝要因と合わせて、疾患発症などの表現型に影響を与える因子の全容を明らかにする。

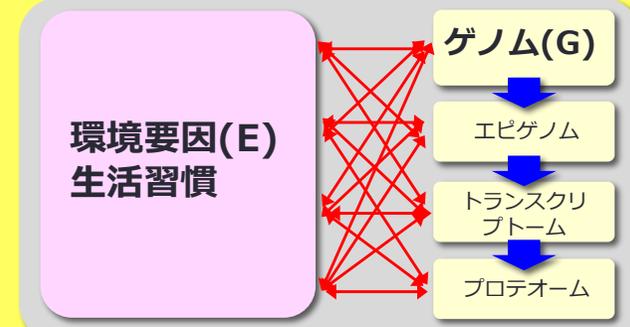
## 研究課題1

メタボロームGWAS解析 (最大5万人)

遺伝要因の解明

代謝物とゲノム情報との関連解析 (GWAS) を行い、各ゲノム多型がタンパク質の機能や量に与える影響（最終的に個人の体質に反映）を代謝物の量的変化から解明する。

メタボロームはすべての要因を反映した、個人の表現型である



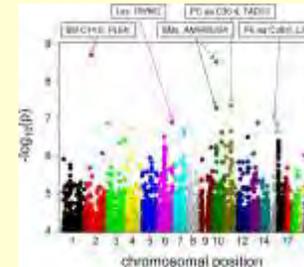
E×Gの全多因子の影響  
(相互作用含む) を反映

メタボローム

# 大規模メタボロームGWAS解析の報告例

- KORA (PLoS Genetics v4 e1000282, 2008)  
284名の参加者から、代謝物と相関のある5種類のSNPを報告
- KORA (Nat. Genetics v42 p137, 2010)  
1809名の参加者から、代謝物と相関のある15種類のSNPを報告
- Framingham Heart Study (Cell Metabolism v18, p130, 2013)  
2076名の参加者から、代謝物と相関のある31種類（23種類は新規）のSNPを報告
- KORA+TwinsUK (Nat. Genetics v46, p543, 2014)  
7824名の参加者（486代謝物+98,346種類の代謝物比）から、代謝物と相関のある145種類（84種類は新規、代謝物の比率との相関も含む）のSNPを報告
- 14 cohort studies (Nat. Commun. V7, 11122, 2016)  
合計24,925名の参加者（共通する123代謝物）から、62種類の相関のあるlocusを報告。

## 最初のメタボロームGWAS解析



284名のメタボローム解析により、代謝に与える5個の遺伝子座を同定

PLoS Genet. v4 e1000282, 2008



## 大規模メタボロームGWAS解析



7824名のメタボローム解析により、代謝に与える137個の遺伝子座を同定

Nat. Genetics v46, p543, 2014

# 症例対照研究（妊婦）の報告例

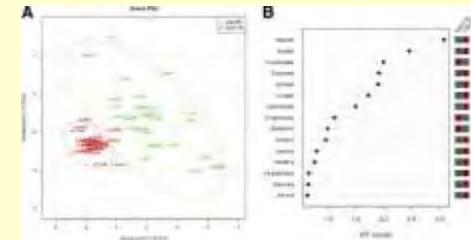
## ■ 妊娠高血圧症候群

- Wayne State Univ.他 (Am. J. Obstet. Gynecol. v58, e1-7, 2013)  
症例対照研究（30名のLate-onset (LO), 30名のEarly-onset (EO) preeclampsiaと59名のコントロール）において血液中の代謝物を比較し、17種類の代謝物で有意な差を同定。
- Norwegian University of Science and Technology他 (Int. J. Mol. Sci. v16, 21520-21538, 2015)  
Prospective Screening Study（妊娠高血圧腎症の中高リスク保有者599名を調査し、実際に妊娠高血圧腎症や妊娠高血圧を発症した対象者（各26名と21名）の血液と尿の代謝物を比較し、有意に差のある代謝物を同定。

## ■ 低出生体重

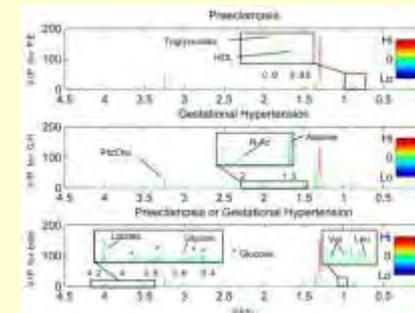
- Medical University of Bialystok他 (Prenat. Diagn. V34, p870-877, 2014)  
コホート研究（770名の妊婦の血液（12~14週）のメタボローム解析を行い、新生児の体重の違いによりリン脂質やビタミン等の濃度が有意に異なることを同定。

## 妊娠高血圧腎症の早期と後期発症妊婦における代謝物プロファイルの比較



Am. J. Obstet. Gynecol. v58 e1-7, 2013

## 妊娠高血圧腎症と妊娠高血圧の各症例における代謝物プロファイル



Int. J. Mol. Sci. v16 21520-21538, 2015

# 生活習慣（環境要因）の影響の解析例

## ■ アルコール

鶴岡 (Environ. Health Prev. Med. v21, p18, 2016)  
2126名のコホート参加者から896名のメタボローム解析によりアルコール摂取量と相関のある代謝物を報告。

## ■ タバコ

KORA (BMC Medicine v11 p60, 2013)  
7341名のコホート参加者から1241名のメタボローム解析を行い、喫煙と相関のある21種類の代謝物を報告

## ■ 食生活（主食）

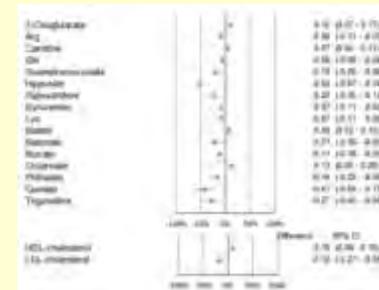
EPIC-Oxford (Am. J. Clin. Nutr. v102, p1518, 2015)  
65,000名の参加者から379名を選択しメタボローム解析を行い、主食（肉、魚、ベジタリアン、等）の違いによる代謝物の変化を報告

## ■ 食生活全般

PLCO Cancer Screening Trial (Am. J. Clin. Nutr. v100, p208, 2014)  
コホート参加者から502名を選択しメタボローム解析

を行い、各種食生活と相関する代謝物を報告

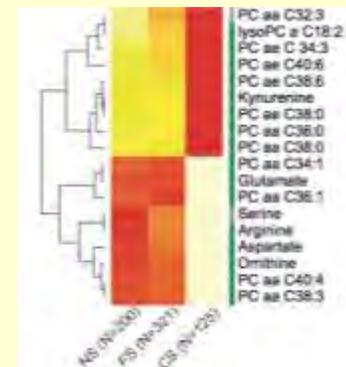
### 飲酒習慣による代謝物変化



アルコール摂取量の違いにより血中のアミノ酸等の量が変化

Environ. Health Prev. Med. v21 p18 2016

### 喫煙習慣による代謝物変化



喫煙により血中の脂質やアミノ酸の量が変化

BMC Medicine v11 p60 2013

# ロードマップ (オミックス解析 7万人)

|  | H28年度           | H29年度             | H30年度             | H31年度             | H32年度             |
|--|-----------------|-------------------|-------------------|-------------------|-------------------|
| <b>GWAS解析<br/>メタボローム解析</b><br>(詳細検査対象者)    | 調整費<br>4,000人   | 多型解析用<br>10,000人  | 多型解析用<br>10,000人  | 多型解析用<br>13,000人  | 多型解析用<br>13,000人  |
| 症例対照 (コントロール含)<br>メタボローム解析<br>(三世代コホート対象者) |                 | 三世代妊婦<br>4,000人   | 三世代妊婦<br>4,000人   | 三世代妊婦<br>2,000人   | 三世代妊婦<br>2,000人   |
| 症例対照<br>メタボローム解析<br>(その他の追跡・2次調査対象者)       |                 | 追跡・二次調査<br>2,000人 | 追跡・二次調査<br>2,000人 | 追跡・二次調査<br>2,000人 | 追跡・二次調査<br>2,000人 |
| <b>合計 (7万人)</b><br>(質量分析は4年間で最大8千人)        | 4,000人          | 16,000人           | 16,000人           | 17,000人           | 17,000人           |
| <b>GRIFIN</b>                              | 生活習慣解析用<br>600人 | 生活習慣解析用<br>600人   | 生活習慣解析用<br>600人   | 生活習慣解析用<br>600人   | 生活習慣解析用<br>600人   |
| <b>不死化リンパ球(GRIFIN)</b>                     |                 | 4年間で2,000検体       |                   |                   |                   |
|  |                 |                   |                   |                   |                   |

GWAS解析5万人の対象は、詳細検査を受けた参加者を対象

宮城県センター型受診者約18,000人、岩手県サテライト受診者約9,000人、三世代センター受診者約11,000人(増加中)、合計約4万人。二次調査受診者から約1万人を選択して、合計で最大5万人

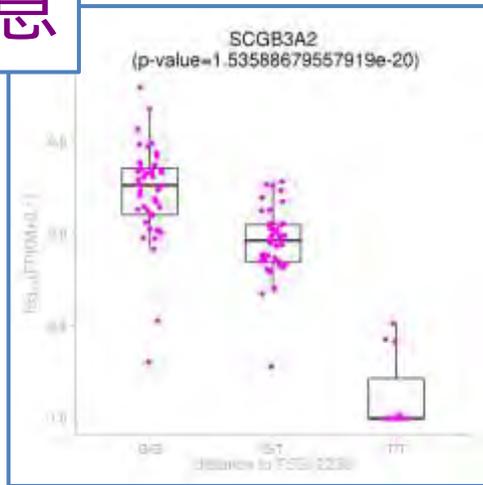
# これまでの成果 期待される成果

## 岩手3層オミックス解析の成果

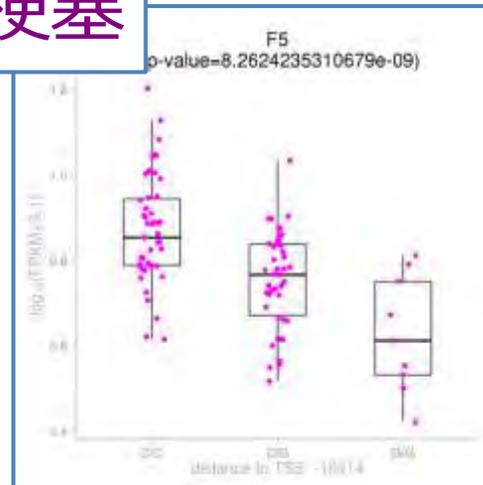
- 100検体の単球、CD4陽性T細胞解析から、それぞれ5万の eQTL を同定

### 疾患との関連が報告された遺伝子の新規 eQTL の例

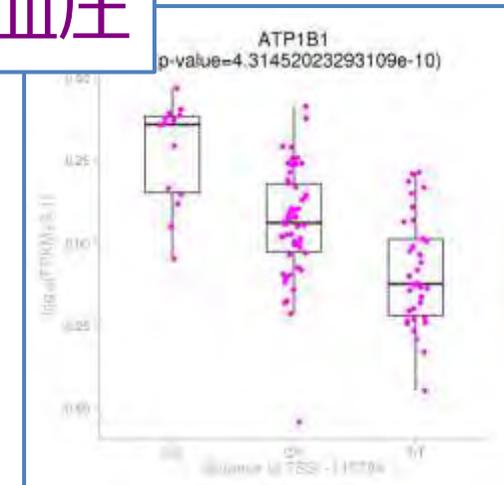
喘息



脳梗塞



高血圧



## 8,000検体の全血・単核球を用いて100万規模の eQTL の同定を目指す

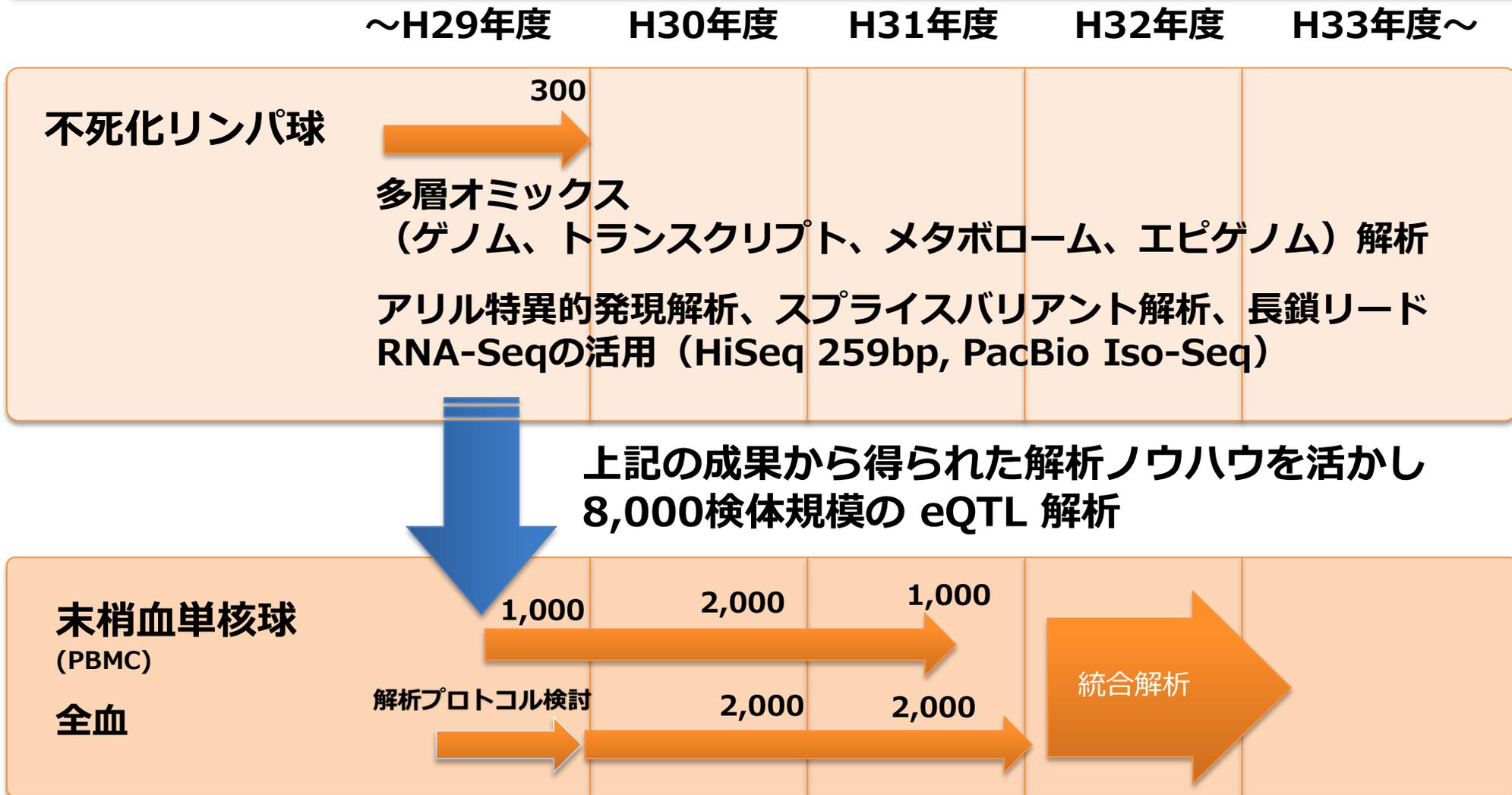
- これまでの多くの eQTL 解析はアレイベース
- RNA-Seqデータを用いたアリル特異的な発現解析による疾患感受性に関わるレアバリエント eQTL の効率的な同定

# TMMが保有する豊富なトランスクリプトーム対象検体

|       | 組織  | 特徴           | 数量    | 保存状態   |
|-------|---|--------------|-------|--|
| ToMMo | 全血<br>(PaxGene)   | ヘテロ集団<br>非株化 | 4,000 | 350検体はWGS済   |
| IMM   | PBMC<br>(BDバキュティナ)  | ヘテロ集団<br>非株化 | 4,013 | 全検体RNA抽出済み   |
| IMM   | B cell<br>CD4+T cell<br>CD8+T cell<br>NK cell<br>Monocyte<br>Neutrophil<br>(FACS) | 均一<br>株化     | 100   | 全検体RNA抽出済み<br>Mono, T cell 発現データあり<br>B cell 発現未解析 |
| ToMMo | 増殖T細胞(IL2+<br>抗CD3)   | 非株化<br>増殖は有限 | 1,351 | 細胞 (凍結) 1KJPN                                      |
| ToMMo | 不死化B細胞<br>(EB virus)  | 株化           | 1,038 | 細胞 (凍結) 1KJPN                                      |

これら検体を有効活用したトランスクリプトーム解析が、  
全ゲノムシーケンスで同定されたSNPの機能アノテーション解析に必須

# トランスクリプトーム解析ロードマップ



疾患感受性に関わるレアバリエーション eQTL を同定し、  
将来の疾患発症リスクを予測、予防医学を目指したゲノム医療に貢献

# 第1段階におけるプロテオーム解析

## ■ フィージビリティ・スタディ（平成25～26年度）

### 質量分析による**血漿プロテオーム**解析プロトコルの確立

#### フェーズ 1

##### 解析技術の確立

- 1) 網羅的解析法 2) 標的定量法

#### フェーズ 2

##### 血液保管条件の検討

- 保管温度・時間の影響

#### フェーズ 3

##### 個人内の生理的変動の確認

- 日内変動（食事の影響）
- 妊娠期間中の変動

#### フェーズ 4

##### ゲノム解析が行なわれたコホート参加者 501人の解析

- プロテオームと健康・医療情報、ゲノム情報との予備的な関連解析
- 疾患バイオマーカー探索の実行可能性を確認
- **Japanese Multi Omics Reference Panel (jMorp)** として、256種の血漿タンパク質の検出頻度と検出ペプチド配列をウェブ公開



#### 貴重な研究基盤として全国で活用

- 疾患バイオマーカーの探索
- 血漿タンパク質の絶対定量法の確立

## 第2段階におけるプロテオーム解析

### ■ 確立した質量分析技術および情報基盤 (jMorp) の活用と展開

- 1) **網羅的解析法** : 網羅的なタンパク質の同定・相対定量 → バイオマーカー探索
  - 1a) **データ依存性解析法** (DDA: Data-dependent Acquisition)  
網羅的な定性解析による**同定**
  - 1b) **データ非依存性解析法** (DIA: Data-independent Acquisition)  
網羅的な**相対定量** (SWATH法など)
- 2) **標的定量法** : 特定タンパク質の**絶対定量** → バイオマーカー候補の検証



1. **コホート検体を用いた共同研究**における、バイオマーカー探索・検証などを目的とするプロテオーム解析。
2. **試料分譲**の要請に対応した、**プロテオーム・データの取得と分譲**。
3. **全ゲノム参照パネルのアノテーション**の充実を目的として、ゲノム解析が実施されたコホート参加者の試料 (**細胞株を含む**) を対象に、プロテオーム解析を含めた**多層オミックス解析**を実施。

**jMorp** :  
血漿タンパク質  
のペプチド情報



# 口腔細菌叢メタゲノム解析による 公開情報基盤の構築（1）

## 1. 概要

- コホート参加者約**2.5**万人から採取した口腔検体（歯垢、唾液）の細菌叢プロファイリングを行い、バイオバンクに格納する
- これら**口腔細菌叢プロファイル**を、個別化予防・医療のためのバイオマーカー探索および介入標的探索のための公開情報基盤として、全国の研究者に提供する

## 2. 背景

- 口腔細菌叢が、歯周病、う歯などの口腔疾患のみならず、糖尿病、動脈硬化症、関節リウマチなど様々な全身性疾患と関連することが報告されている

## 3. 作業仮説

- 口腔細菌叢と口腔疾患及び全身性疾患との関連を、前向きコホート調査により縦断的に解析することにより、特定の細菌種が疾患の危険因子あるいは逆に抑制因子となるか否かが明らかになる可能性がある
- 口腔細菌叢をモニターすることにより、これら疾患のリスク判定や予防に活用できる可能性がある
- 口腔細菌叢が、これらの疾患の予防・治療のための介入標的として利用できる可能性がある

# 口腔細菌叢メタゲノム解析による 公開情報基盤の構築（2）

## 4. 検体収集の現状

- **対象者**： 地域住民コホート及び三世代コホート調査において、宮城県内の7ヶ所の地域支援センターで歯科検診を受診した参加者。

| 地域住民コホート | 三世代コホート | 合計      |
|----------|---------|---------|
| 約 1.7万人  | 約 0.8万人 | 約 2.5万人 |

- **検体**： 唾液（～3 mL）、歯垢（大臼歯より）
- **収集期間**： 2013年～2016年5月（ベースライン調査での検体収集は終了）

## 5. 解析手法：フィージビリティ・スタディにより検討

### ■ 16Sリボソーム遺伝子解析

- 可変領域のPCR産物の標的シーケンス解析による細菌叢プロファイリング。
- MiSeqを用いた多検体同時並列シーケンス解析により、簡便、安価。

### ■ 全ゲノムショットガン解析

- PCRフリーのため低バイアス・高精度。微量細菌を含め、幅広い菌種を高解像度で同定・定量可能。
- 遺伝子機能に基づくプロファイリングが可能。

# ゲノム・オミックス解析

## 第2段階計画概要

---

### 1. はじめに

### 2. 大規模ゲノム解析の意義

- 1) 海外のコホート・バイオバンクにおけるゲノム医科学研究の現状、大規模アレイ解析との比較
- 2) 疾患関連遺伝子研究に向けた基盤の提供
- 3) 大規模ケース・コントロールスタディでの活用

### 3. ゲノム解析の実施計画

- 1) ゲノム解析のロードマップ
- 2) 大規模SNPアレイ解析と全ゲノム情報復元
- 3) 全ゲノムリファレンスパネルの拡充
- 4) 全ゲノム解析の手法とその選択

### 4. オミックス解析等の実施計画

- 1) 多層オミックス参照パネルの構築と展開
- 2) 大規模メタボローム解析
- 3) エピゲノム解析
- 4) その他のオミックス解析等

### 5. 個別化予防に向けた統合解析のための基盤構築

- 1) 発症リスク予測手法開発
- 2) 統合・知識データベースと病院連携

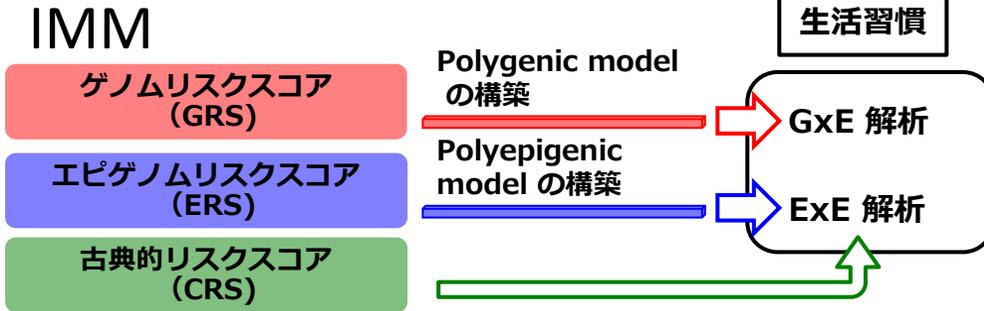
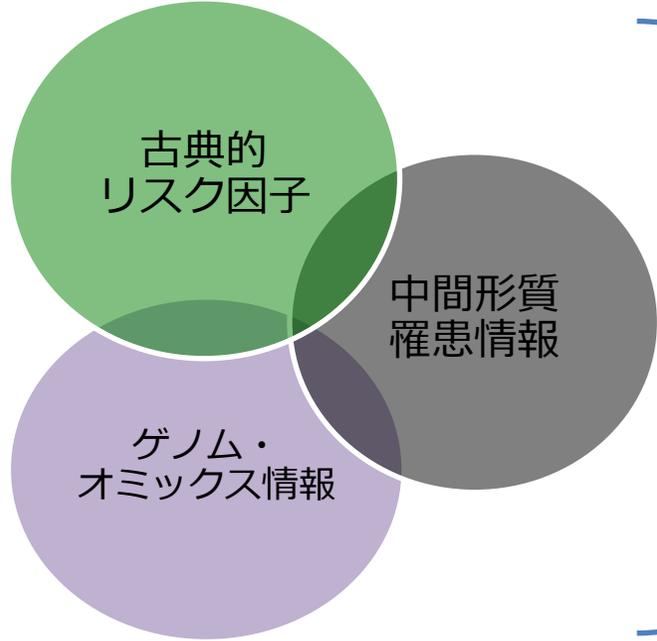
### 6. 遺伝情報回付

### 7. ゲノム医療実現のための人材育成

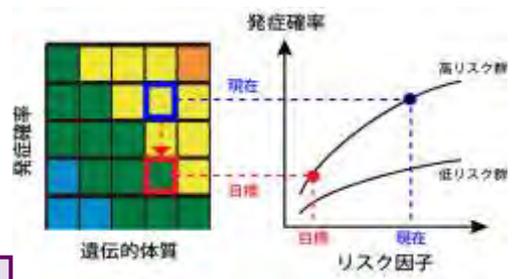
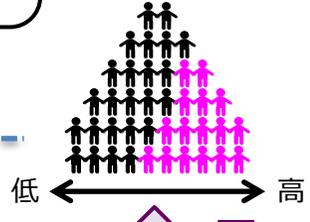
### 8. まとめ

# ゲノム情報、オミックス情報、中間形質、古典的リスク因子を考慮した疾患発症予測モデルの構築と検証

## メガバンク情報の拡充



## ToMMo



発症リスク予測手法の開発

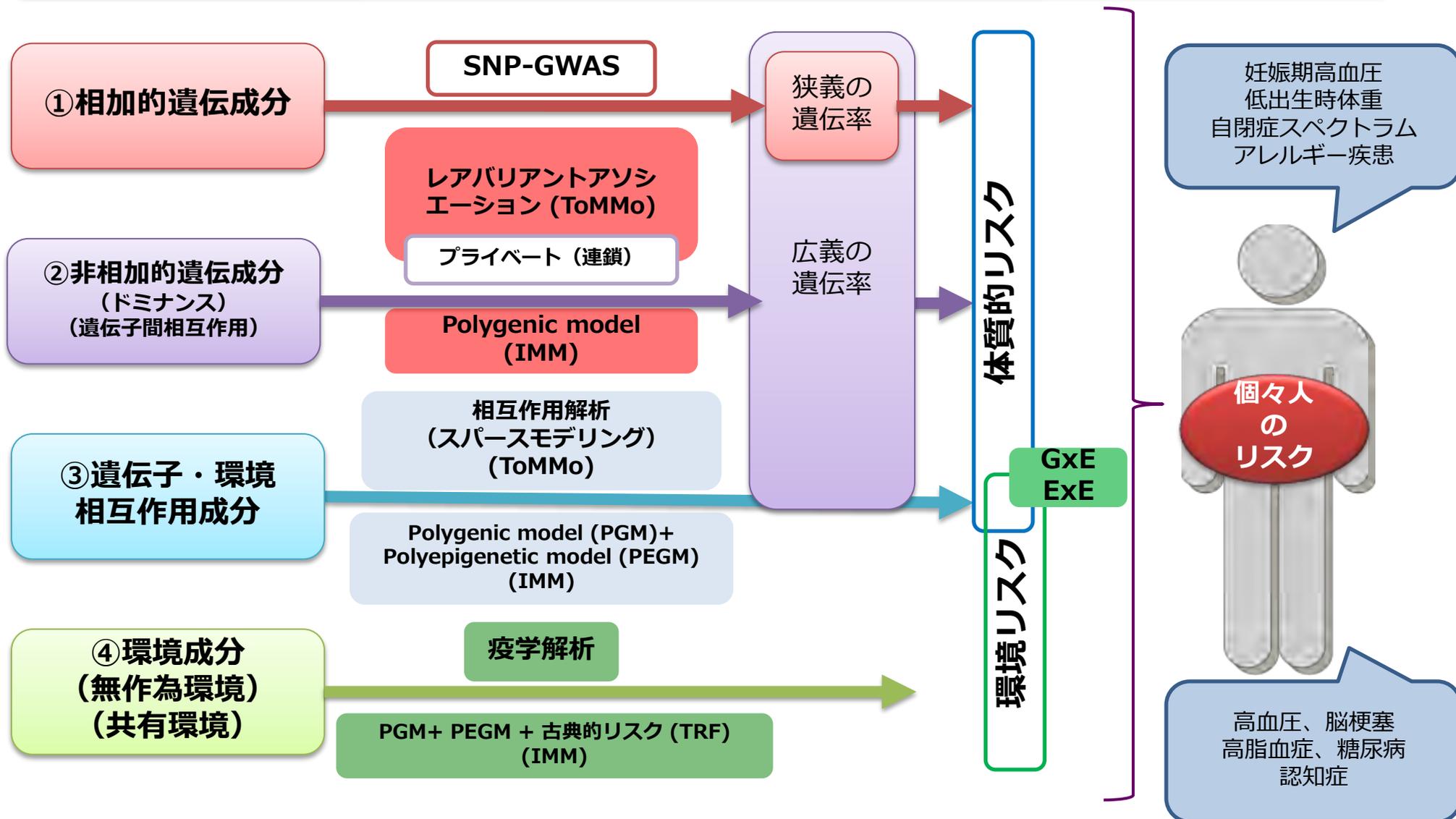
発症リスク予測手法の精度検証

コホート連携による検証も視野に



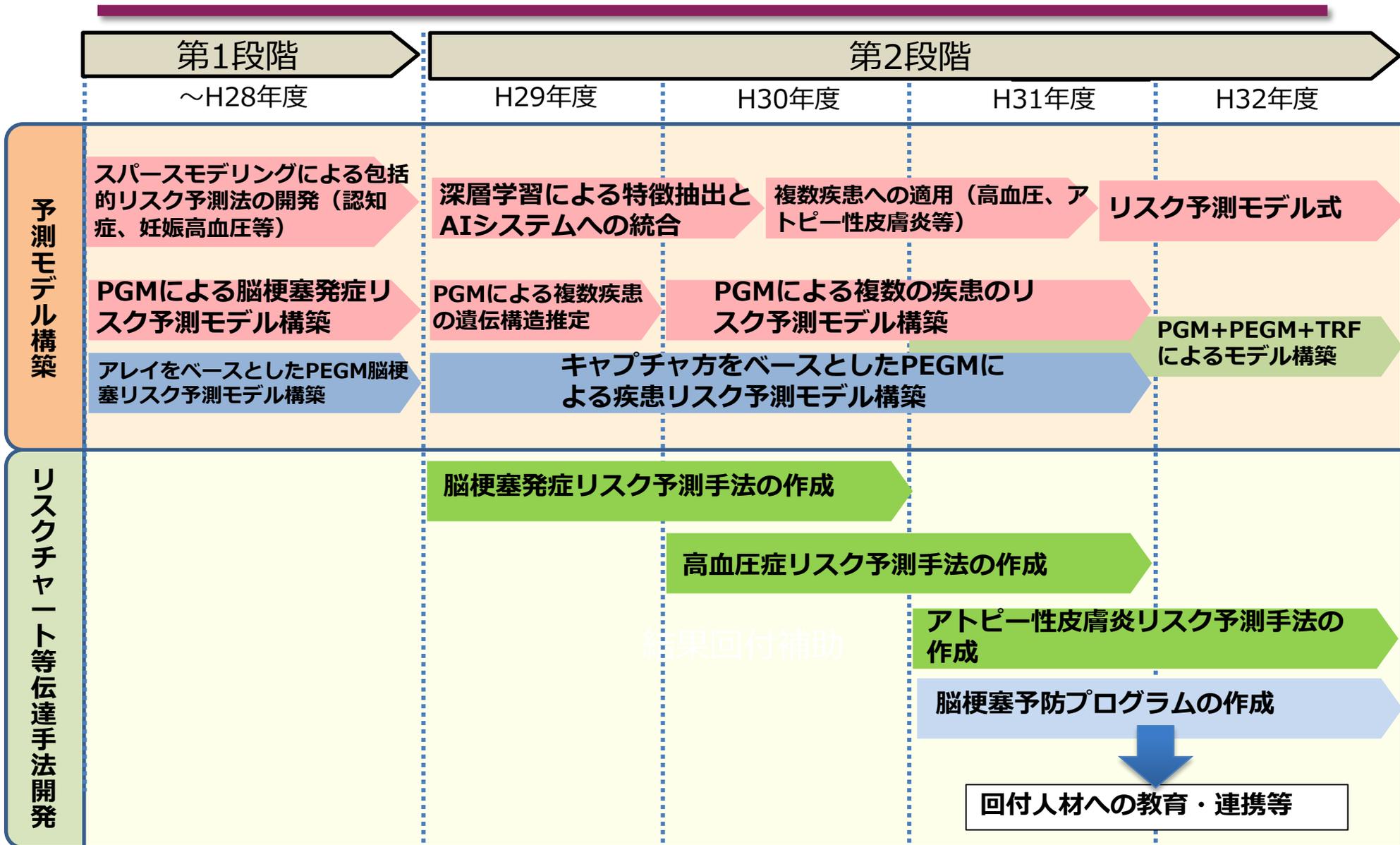
被災地における発症リスク予測手法に基づく次世代型予防医療の実現

# 疾患・健康状態に関する個々人の包括的リスク予測



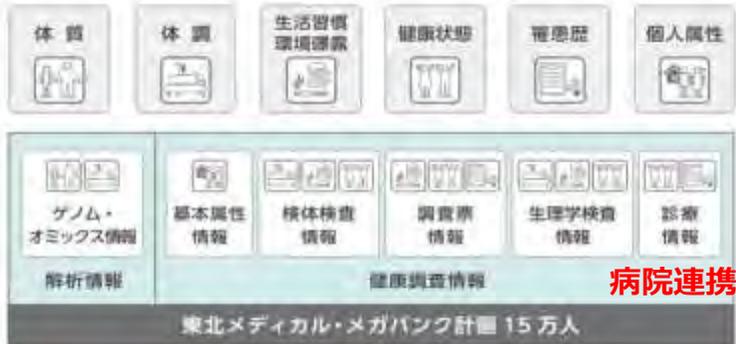
従来のゲノム医学でのリスク予測は①のみを不正確に扱うしかなかった。  
TMMでは①、②、③、④を適切に取り込んだリスク予測を行う。

# 疾患発症リスク予測モデルのロードマップ



# 統合・知識データベースの進捗と今後の計画

## 15万人ゲノムコホートデータ

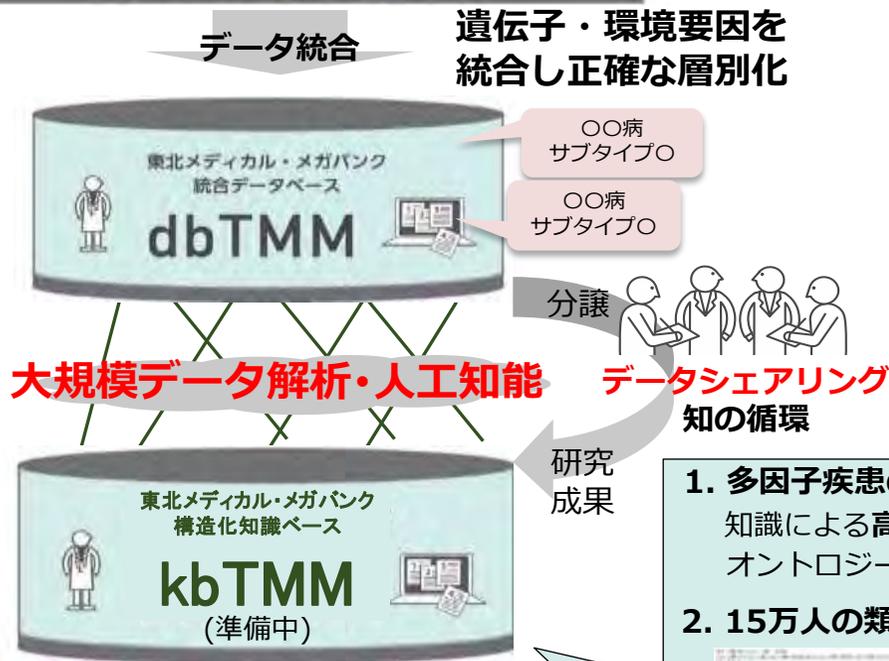


## 第1期の成果

- ✓ 健康調査情報及び解析情報を統合する、統合データベースdbTMMを開発
- ✓ 1,070人分の全ゲノム配列情報、健康調査情報を制限付公開、全国の研究者が手続きを経て利用中
- ✓ ゲノム・オミックス解析情報及び健康調査情報を統合した精確な層別化
- ✓ ゲノム医療実現推進のデータシェアリングの基盤

## 第2段階での計画

- ✓ 二次調査、追跡調査による収載する情報の種類・量の拡大、時系列データ統合・機能強化
- ✓ 病院連携による診療情報の集積
- ✓ 情報を統合したフェノタイピング
- ✓ バイオバンク連携としてインターネット公開版統合データベースdbTMM-pubの開発、横断検索への展開
- ✓ 大規模データ解析・人工知能による構造化知識ベースkbTMMの開発



15万人の類似症例のクラスタリング  
ゲノム医療・研究の知識ベース

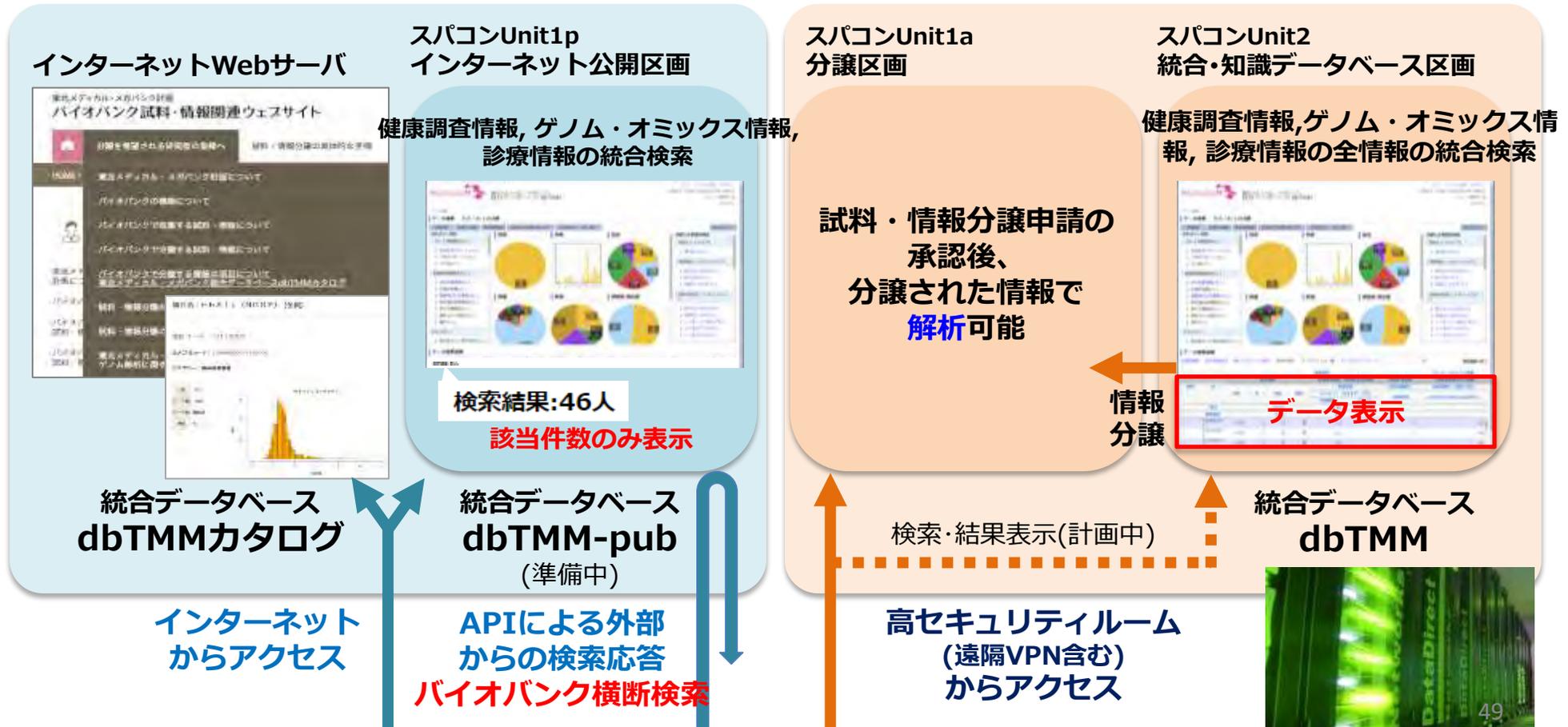
1. 多因子疾患の遺伝要因・環境要因の知識ベース  
知識による高度な利用、精確なフェノタイピング支援  
オントロジーに基づく相互運用性(RDF)、遺伝情報回付等のゲノム医療支援
2. 15万人の類似症例のクラスタリング機能



- 参加者個々の超高次元ベクトルの変数選択とクラスタリング
- 類似症例検索による統合データベースdbTMMの高度な層別化支援、疾患発症予測への応用

# 統合データベースdbTMM・dbTMM-pub・dbTMMカタログ

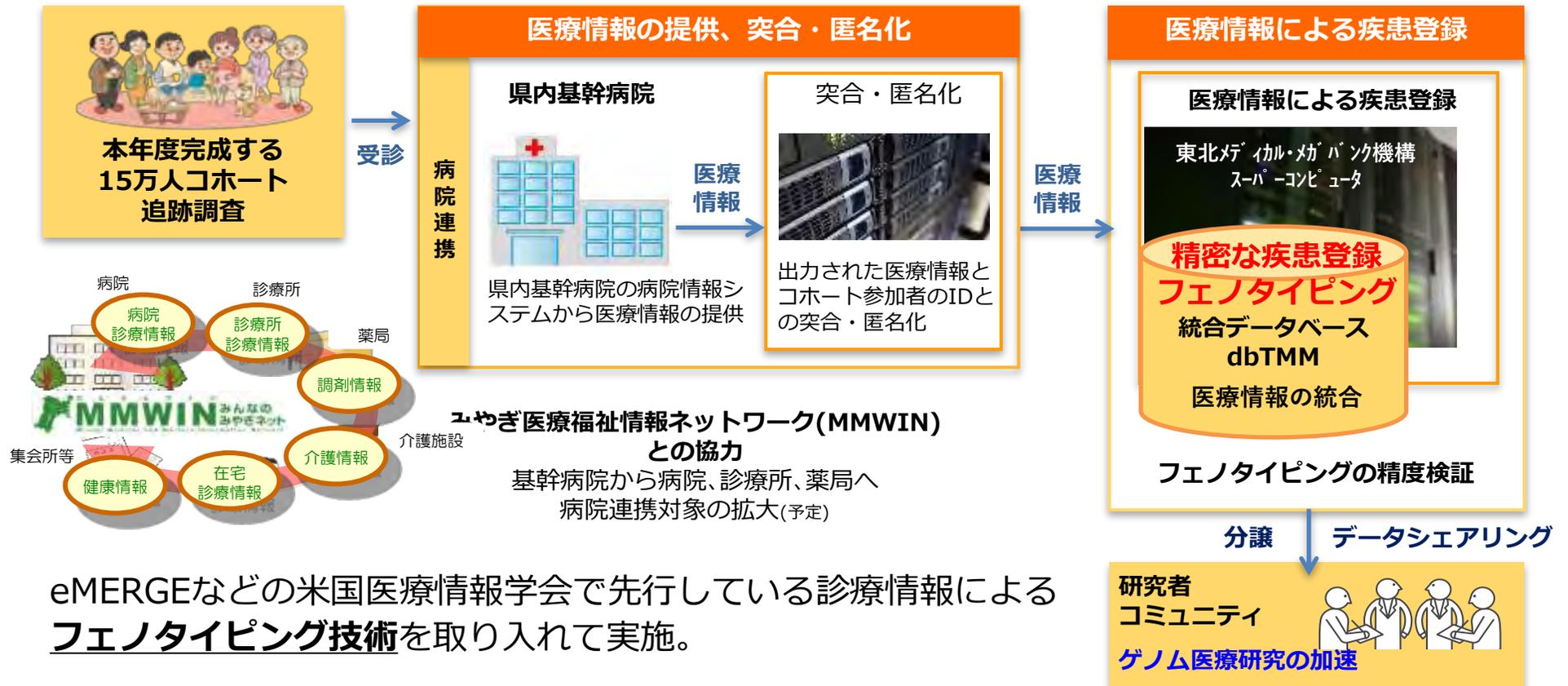
収集した健康調査情報、ゲノム・オミックス情報、診療情報(予定)を統合した、**統合データベースdbTMM**、インターネットから利用可能な**dbTMM-pub**および**カタログ**を提供し、**データシェアリングの促進によりゲノム医療実現のための解析研究を支援する**



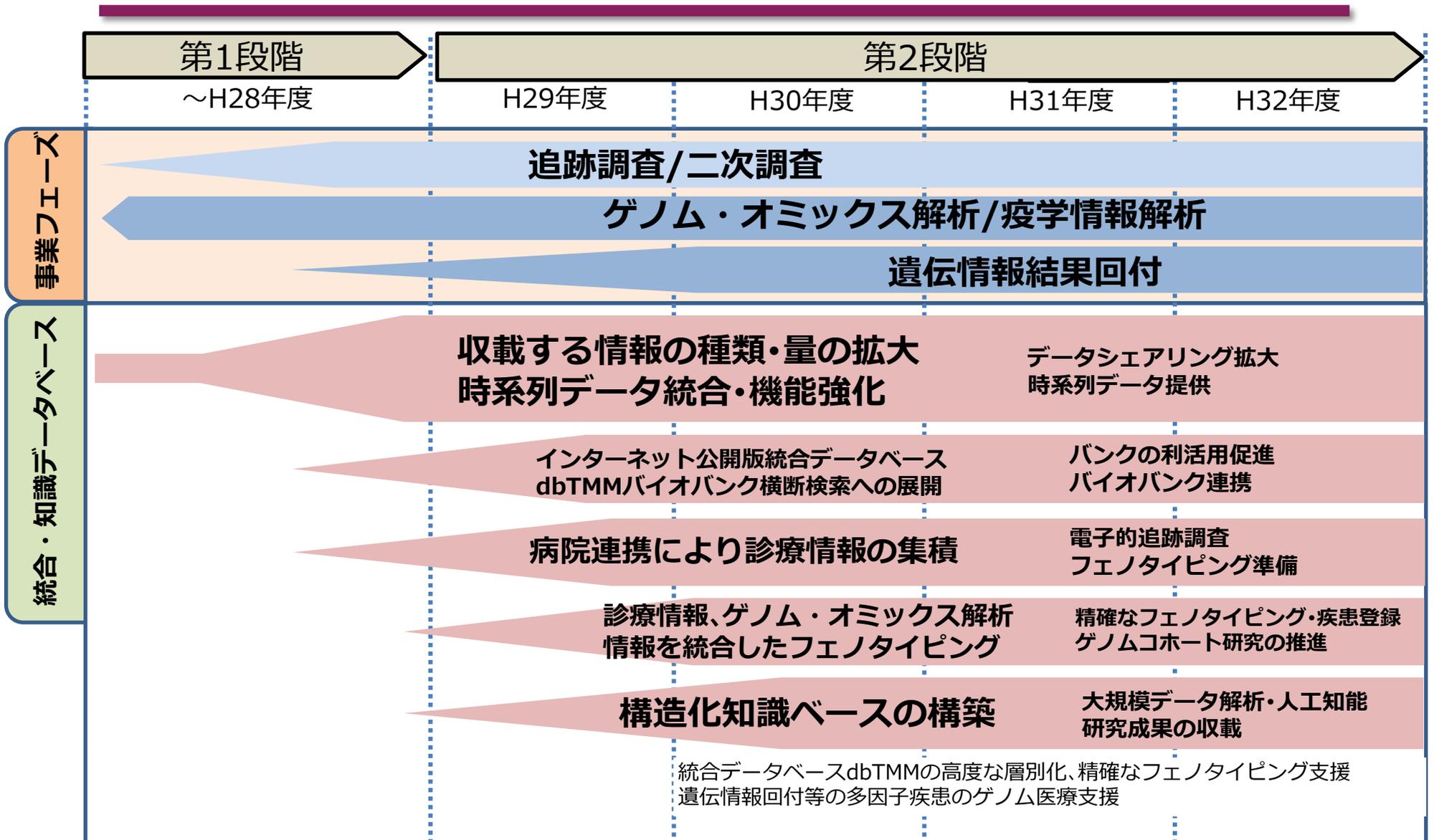
# 病院連携による医療情報の利活用による追跡調査

- 県内基幹病院と連携し、コホート参加者の医療情報を収集する
- 医療情報、ゲノム・オミックス情報、生活習慣・環境暴露などの健康調査情報と統合した疾患登録、フェノタイプングを実施し、**全国のゲノム医療研究へデータシェアリング**する

※ただし、診療情報はそのままデータシェアリングせず、ゲノム医療研究に必要な解析に供する情報に整理しての提供を予定している。



# 統合・知識データベースのロードマップ

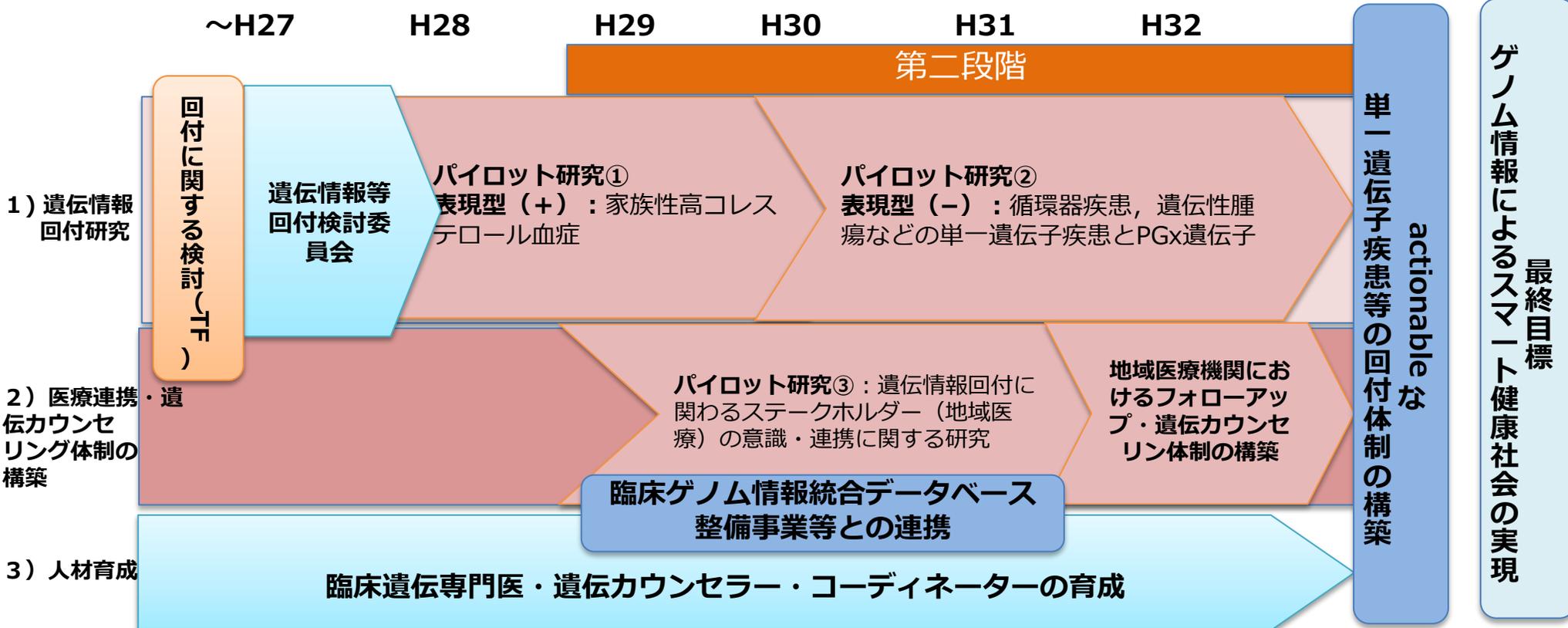


# 遺伝情報回付の取り組み（1）（案）

## ゲノム情報を利用した個別化医療・個別化予防のために

- 東北メディカル・メガバンク計画では、当初より個人への遺伝情報の回付を想定した説明と同意を取得しているコホート研究である
- 平成25年度のリクルート開始以後、遺伝情報回付検討タスクフォースの立ち上げ、個人への回付に向けての検討を開始
- コホート研究での個人への遺伝情報等回付は世界的にも例が少なく、様々課題点が抽出された。
- 重層的なパイロット研究計画をおこないながら、上記の実現を目指す

### 単一遺伝性疾患等（リスクが明らかで浸透率が高い疾患）

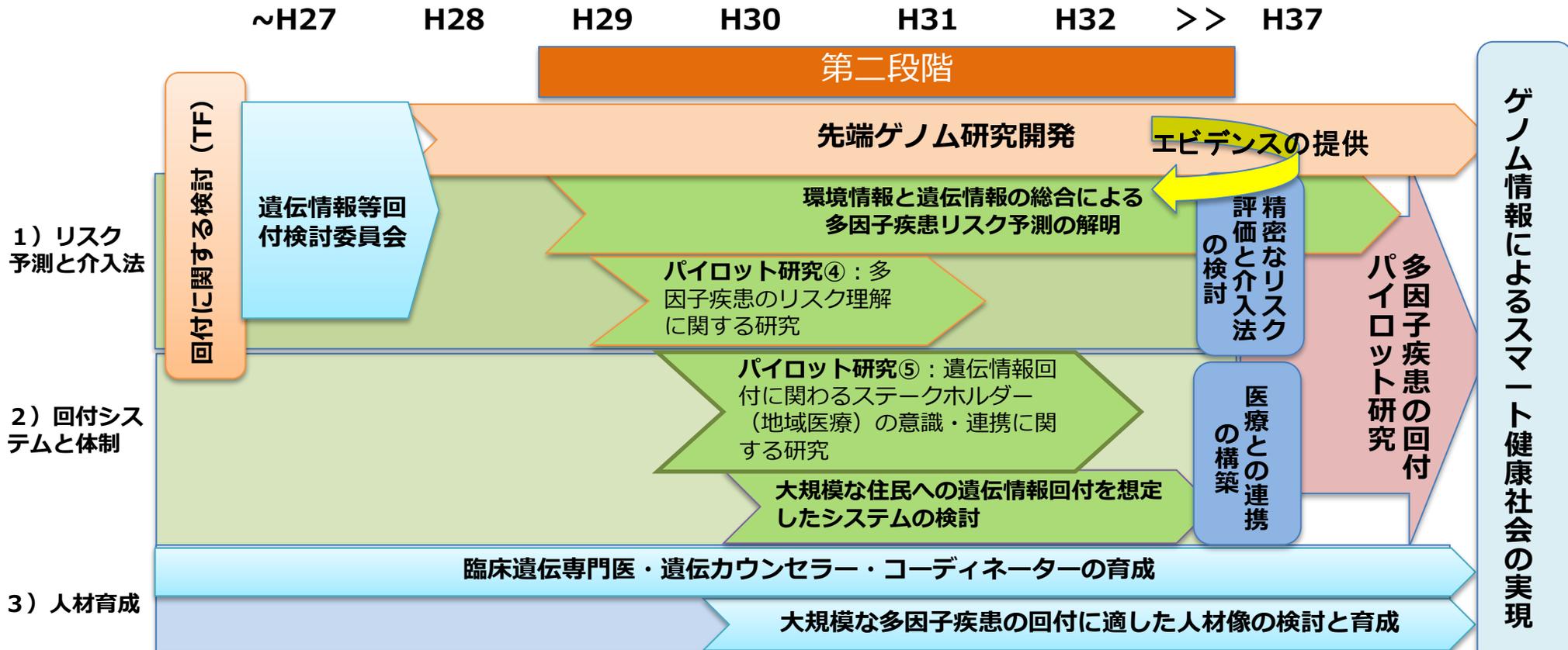


# 遺伝情報回付の取り組み（2）（案）

## ゲノム情報を利用した個別化医療・個別化予防のために

- リスク評価は難しい多因子疾患に関しては、「リスク評価+介入法」の回付に向けての基盤整備を行う
- ゲノムコホートの追跡調査とゲノム情報の統合による精緻なリスク評価と介入法の検討
- 単一遺伝子疾患等のパイロット研究の成果を踏まえた、回付後の地域医療、保健システムと協働したフォローアップ体制の整備
- ゲノム医療実現推進プラットフォーム事業（先端ゲノム研究開発）とも連携

### 多因子疾患（リスク評価が難しい疾患）



# 第1段階におけるToMMoの人材育成の実績

## ゲノム・メディカルリサーチコーディネーター (GMRC)

コホート調査における参加者リクルート、地域医療の支援を行う人材。看護師、保健師、検査技師、一般適性者を被災地域で採用し、採用後、必要な教育、研修を実施して試験に合格したものをGMRCと認定。**現在、199名のToMMoGMRC認定者が在職。(実働数155名)**  
(全国の人類遺伝学会認定GMRC482人)

## 認定遺伝カウンセラー (CGC) / 臨床遺伝専門医

コホート参加者に対するゲノム情報の説明、次世代ゲノム医療・個別化医療を担う人材。遺伝カウンセラーコース（平成25年度より東北大学医学系研究科公衆衛生学専攻修士課程に開設）には3名在籍中。**平成27年度卒業生2名**。平成26年度に1名CGCを採用。  
メガバンク所属の臨床遺伝専門医は現在6名。一期中に3名が新規認定を受けた。

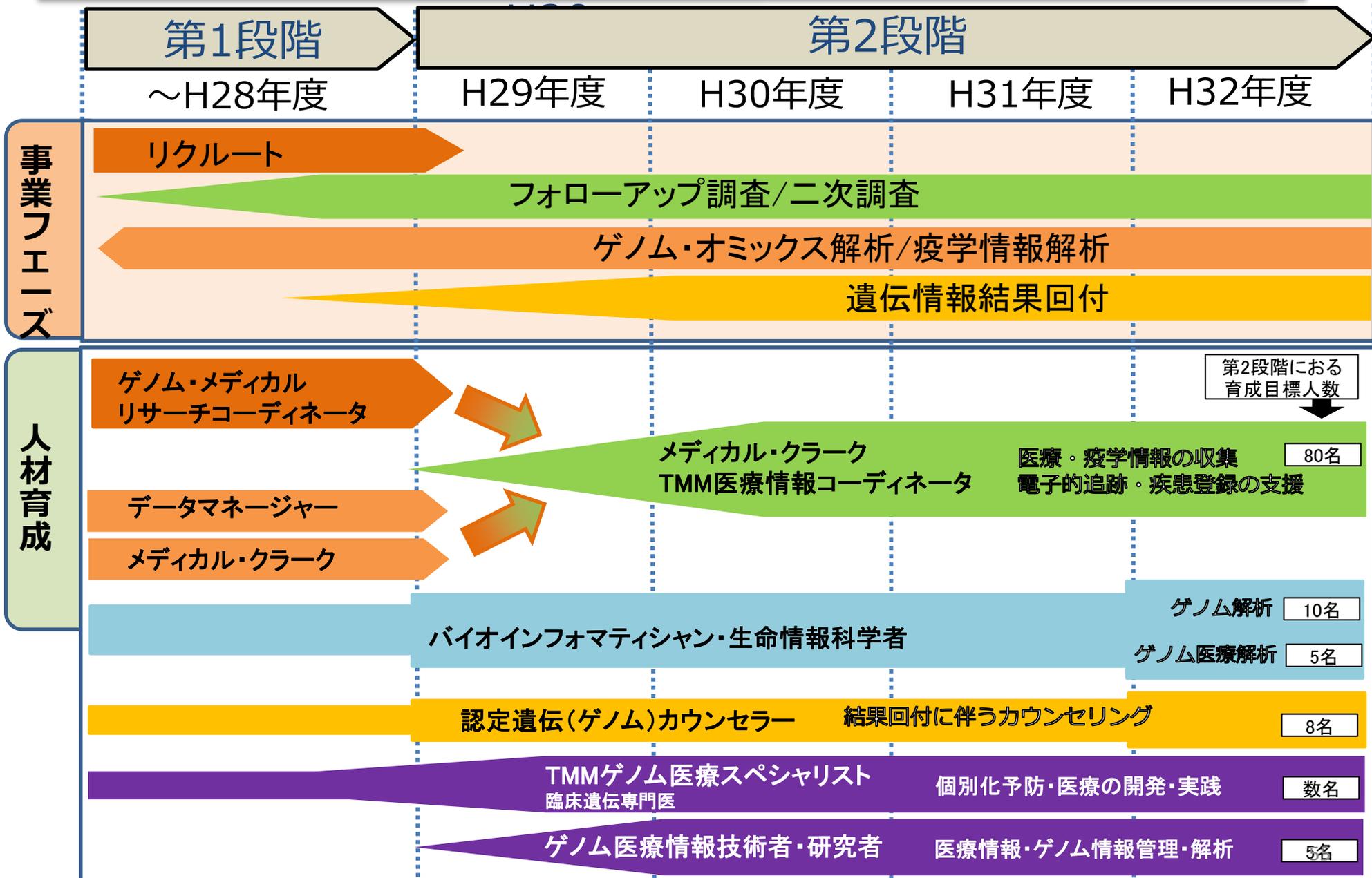
## データマネージャー (DM) / メディカルクラーク (MI) / 医療情報技術者 (MC)

コホート情報、医療情報のマネジメントを行う人材。On the job trainingで養成している。現在、DM 9名、MC 1名、MI 2名が従事している。

## 生命情報科学者 / バイオインフォマティシャン

ゲノム情報のインシリコ（コンピュータ）解析に従事する人材。  
現在、教授・准教授3名、助教8名、助手・技術職員12名体制となっている。  
平成24年度より、大学院医学系研究科、情報科学研究科に専門の授業を開講した。  
**平成27年度は15名（4名社会人博士課程含む）の修士・博士課程の学生の教育を担当した。**

# 第2段階のゲノムコホート研究必要な人材育成ロードマップ



# ゲノムコホート研究にかかわる職種(参考)

□東北メディカル・メガバンク計画の推進のために育成、確保が必要な人材配置マップ

