

# Database of Pathogenic Variants (DPV)の開設

慶應義塾大学医学部臨床遺伝学センター

鈴木寿人、上原朋子、武内俊樹、小崎健次郎

## はじめに

次世代シーケンサーの登場以降、網羅的遺伝子解析は臨床医にとっても身近な診断手法となりつつある。得られた大量のバリエーションの解釈にはさまざまなデータベースが必須である。なかでも病的バリエーションを収集したデータベースとの照合は診断に直結することも多く、非常に有用である。

## 国内外のデータベースの現状

|             | 国内                                | 海外  |
|-------------|-----------------------------------|---|
| 疾患特異的データベース | RAPID(免疫不全症) 難病研究班DB              | CFTR(嚢胞性線維腫) PAHvdb (PAH) など                    |
| 網羅的疾患データベース | iJGVD (一般集団) HGVD (一般集団) HGV (疾患) | 1000G (一般集団) ExAC (一般集団) ClinVar (疾患) HGMD (疾患) |

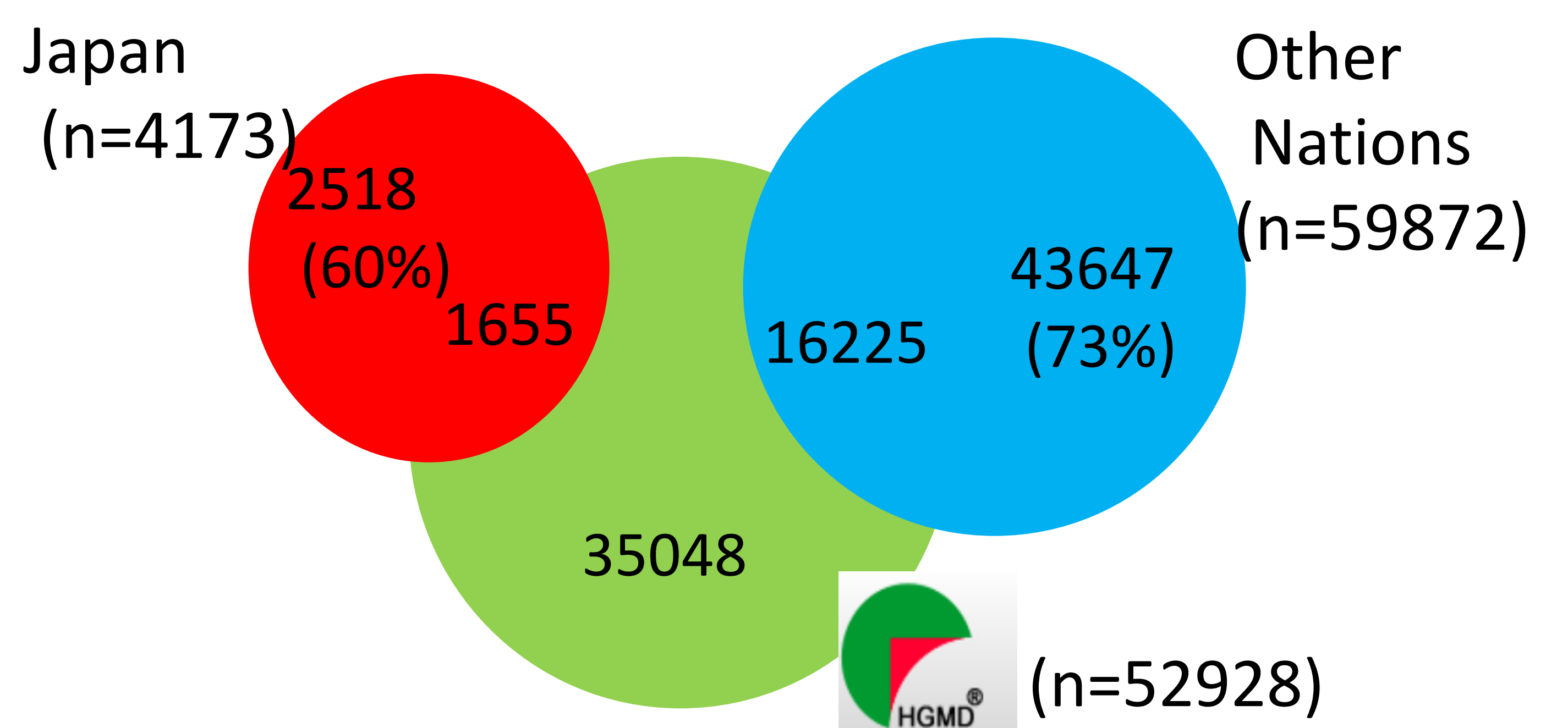
国内は各難病研究班を中心とした非公開・限定公開のDBは多数存在すると考えられる。疾患を網羅的に収集するHuman Genome Variationもあるが、論文の前向き登録になっている。

国外のClinVarやHGMDに相当する、これまでの研究成果を含めて収集する網羅的な疾患バリエーションDBは国内にはなかった。

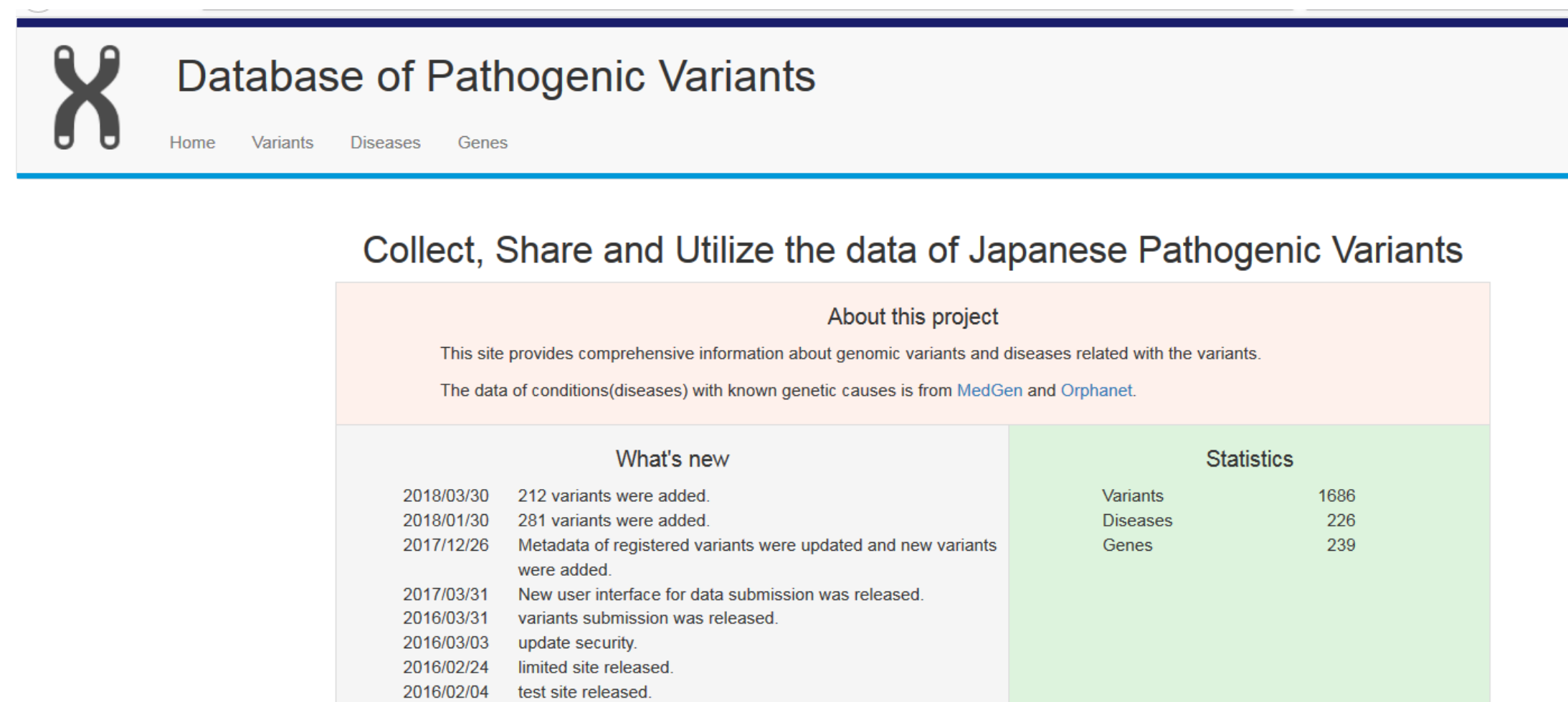
## 日本発論文のデータベース掲載状況

PUBMEDで“mutation” “NOT cancer” をキーワードに検索し、執筆者の所属で分類を行った。HGMDに登録のある文献IDとの比較を実施した(右図)。

日本発の論文のうちDBに利用されているものは約40%であり、残りはDBに未登録あるいは主論文となっていない状況であった。



## DPVの開設



<http://dnv.cmg.med.keio.ac.jp/dpv-pub/top>  
検索: DPV variant

### 登録要件

- ①サンガー法により、変異が確認されていること
- ②日本人の健常人データベースにおいて多型頻度が3%未満であり、十分に低いこと
- ③ピアレビュー付の論文でアクセプトされていること

登録後、ゲノム座標、頻度情報をキュレーションしメタデータを付記して掲載している。

2018年9月までに2238件のバリエーションの登録を行った。今後、ペースアップしていく予定

## マニュアルキュレーションの必要性

- HGVS表記に準拠しない変異の記載/誤記載
- 参照配列の未記載
- 正常バリエーションと疾患バリエーションと混在した論文からのバリエーション抽出
- 欧米人の頻度情報のみからの病原性の判断

## DPVの今後の方針/予定

- キュレーションの人材募集と育成
- 小児科学会を通して各分野の専門家との協力・連携することが決まった。今後、論文データあるいは未発表であっても公開可能なデータの収集を行う。

## まとめ

データベースはその掲載されるデータ量と質によって評価される。今後、小児科学会をはじめ、その他の学会を通じて多くの研究者より提供可能なデータを募り、データベースを充実させていきたい。



Licensed under a Creative Commons表示4.0国際ライセンス (c)2018 鈴木寿人 (慶應義塾大学医学部)