

# PubCaseFinder : 稀少・遺伝性疾患検索プラットフォームの構築

藤原豊史<sup>1)</sup> 山本泰智<sup>1)</sup> 金進東<sup>1)</sup> 山口敦子<sup>1)</sup> 申在紋<sup>1)</sup> 宮冬樹<sup>2)</sup> 武藤勇<sup>3)</sup> 山本利也<sup>3)</sup> 浅野由衣<sup>4)</sup> 高木利久<sup>5)</sup>

fujitoyo

1) ライフサイエンス統合データベースセンター 2) 東京医科歯科大学 3) BITS 4) PENQE 5) 富山国際大学

fujiwara@dbcls.rois.ac.jp

我々が開発する稀少・遺伝性疾患検索プラットフォーム「PubCaseFinder」は、患者の症状を入力として（日本語・英語に対応）、約7000件の遺伝性疾患、約4000件の稀少疾患、約30万件の症例報告を、患者と関連性の高い順にランキングとして提示するため、鑑別診断に活用できる。また、エクソーム解析などのゲノム解析で検出された多数のバリエーションから、対象疾患の原因変異のバリエーションを絞り込む場合にも、患者の症状を利用することができる。本発表では、論文発表済みの17検体を用いて、既存ツールPhenIXとバリエーション絞り込み機能の精度を比較したので報告する。昨年以降のアップデートとしては、PubCaseFinderに組み込まれている症状検索ツールPhenoTouchの改良を行った。PubCaseFinderが入力として利用するHuman Phenotype Ontology (HPO) は、約13,000件の用語を含み、その中から適切な用語を探すのは困難な場合が多い。PhenoTouchは効率よく目的のHPO用語を検索することができる。また、PubCaseFinderのRDF化を行い、SPARQLを用いた情報取得が可能になったので、その利用例を紹介する。

**患者の症状を複数入力 (各症状の詳細情報を表示可能)**

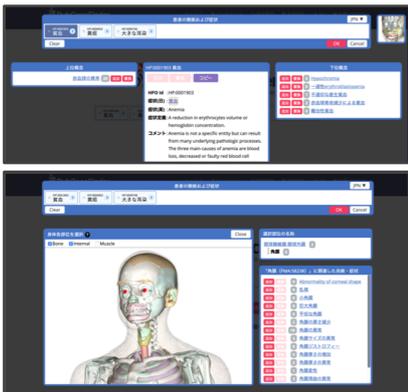
**患者の症状セットとの関連度順に疾患がランキング提示される**

**入力した遺伝子に関連する疾患のみのランキングが提示される**

**日本語・英語で入力可能 (同義語も合わせて検索)**

## PhenoTouchについて

オントロジー階層の可視化や、ヒト3Dモデルでの症状選択により、約13,000件の症状セットの中から目的の症状を探しやすくするツールを開発。



## PubCaseFinder RDFについて

鑑別診断、遺伝学的検査に有用な、疾患-症状関連データ、疾患-遺伝子関連データ、遺伝子-症状関連データを中心に、PubCaseFinderで利用しているデータをRDF化し、SPARQLエンドポイントを公開した。

詳細な情報は以下のポータルサイトを参照ください。

<https://pubcasefinder.dbcls.jp/portal>

### 統計情報

データ名	データ数
疾患-症状関連データ	180,410
疾患-遺伝子関連データ	13,800
遺伝子-症状関連データ	201,026

## PubCaseFinderのバリエーション評価精度

東京医科歯科大学の宮冬樹先生のところで、疾患遺伝子変異解析の専門家と臨床医が候補遺伝子を絞り込むことでpathogenicなバリエーションを同定した17検体（論文で報告済み）を用いて、PubCaseFinderとPhenIXの精度を比較した。

### 評価方法

**【入力】**  
 ・ 症状：HP:0001250,発作, HP:000648,視神経萎縮  
 ・ バリエーション：vcfファイル  
 →PubCaseFinderはバリエーションを遺伝子リストに変換して入力 (DHODH, SF3B4, ...)



**【出力】**  
 ・ 疾患ランキング (PubCaseFinder)  
 ・ 遺伝子ランキング (PhenIX)  
 論文報告遺伝子の順位を確認

性別	年齢	臨床所見	臨床所見(HPO)	解析方法	解析/バリエーション後の候補バリエーション数	論文で報告された疾患関連遺伝子	報告論文	PubCaseFinder解析結果		PhenIX解析結果	
								ランキング	論文報告遺伝子の順位	ランキング	論文報告遺伝子の順位
男	3歳	小脳低形成	HP:000121,小脳低形成	エクソーム解析 ( Trio )	34	NRXNPH2	Bain JM, et al. Am. J. Hum. Genet. 99, 3, 728-734 (2016).	24	1	17	10
男	13歳	てんかん, 視神経萎縮	HP:0001250,発作 HP:000648,視神経萎縮	エクソーム解析 ( Trio )	26	RTN4IP2	Okamoto et al. J. Hum. Genet. 62, 927-929 (2017).	18	1	11	10
男	10歳	皮膚乾燥症	HP:000973,乾燥性皮膚病	エクソーム解析 ( Trio )	27	ALDH18A1	Nozaki et al. Brain Dev. 38, 678-684 (2016).	23	1	8	1
男	5歳	白質異常, 皮質異常	HP:000500,大脳白質の異常 HP:010070,Cortical visual impairment	エクソーム解析 ( Trio )	36	KIF1A	Okamoto N, et al. J. Hum. Genet. 59, 639-641 (2014).	25	1	20	4
男	1歳	精神発達遅滞, 骨質異常	HP:000432,低身長 HP:000924,骨質の異常	エクソーム解析 ( Trio )	16	FAM42C	Tamai et al. Am. J. Med. Genet. A 176, 682-686 (2018).	7	1	7	2
女	1歳11か月	macrocephaly, cortical dysplasia	HP:000256,Macrocephaly HP:000259,Cortical dysplasia	エクソーム解析 ( Trio )	36	AKT3	Negishi Y, et al. BMC Med. Genet. 18, 4 (2017).	33	1	20	2
女	2歳	小脳低形成, 脳幹低形成	HP:000121,小脳低形成 HP:000265,脳幹低形成	エクソーム解析 ( Trio )	18	TUBA1A	Yokoi S, et al. Sci. Rep. 5, 15165 (2015).	15	1	8	1
女	2歳	小脳症, 脳室周囲性石灰白質病, 肥満症	HP:000252,小脳 HP:000765,脳室周囲性石灰白質病 HP:000568,小脳症	エクソーム解析 ( Trio )	21	PLK4	Tsutsumi M, et al. Eur. J. Hum. Genet. 24, 1702-1706 (2016).	7	1	9	3
男	4歳	半球間葉細胞, 脳梁低形成	HP:000209,脳梁低形成	エクソーム解析 ( Trio )	34	NFIA	Negishi Y, et al. Hum. Genome Var. 2, 15007 (2015).	13	1	20	18
女	NA	脳梁欠損, 知的障害	HP:000274,脳梁無発生 HP:000124,知的障害	エクソーム解析 ( Trio )	3	EPGS	Hori et al. Sci. Rep. 14, 3552 (2017).	1	1	1	1
女	NA	多小脳回, 異所性石灰白質	HP:000126,多小脳回 HP:000281,異所性石灰白質	エクソーム解析 ( Trio )	23	NEDD4L	Kato et al. J. Hum. Genet. 62, 861-863 (2017).	12	1	13	8
女	NA	小脳症, 低身長, 自閉症	HP:000252,小脳 HP:000432,低身長 HP:0000717,自閉症	エクソーム解析 ( Trio )	16	ASXL3	Hori et al. J. Hum. Genet. A 170, 1863-1867 (2016).	6	2	10	8
女	NA	小脳症, てんかん, 自閉症	HP:000252,小脳 HP:000121,小脳低形成 HP:000124,知的障害	エクソーム解析 ( Trio )	46	STAMBP	Hori et al. J. Hum. Genet. 63, 957-963 (2018).	15	1	20	2
男	NA	巨脳症	HP:0001355,巨大脳症	エクソーム解析 ( Trio )	34	MVNC	Kato et al. J. Med. Genet. 56, 388-395 (2019).	40	1	15	3
女	NA	小脳症	HP:000252,小脳	エクソーム解析 ( Trio )	5	ASPM	F. Miya et al. Sci. Rep. 5, 9331 (2015).	3	1	2	1
女	NA	脳梁欠損	HP:0001274,脳梁無発生	エクソーム解析 ( Trio )	38	EPGS	Hori et al. Sci. Rep. 14, 3552 (2017).	15	1	15	1
男	NA	水頭症	HP:000238,水頭症	エクソーム解析 ( Trio )	25	TUBA1A	Yokoi S, et al. Sci. Rep. 5, 15165 (2015).	8	1	12	1

結論：類似システムの中で最も精度が高いと報告 (Pengelly et al. 2017) されている PhenIXよりも、PubCaseFinderの精度が高い結果となった。PubCaseFinderで唯一2位となった検体において、論文から患者の詳細な症状を抽出し、再解析を行った結果、1位となり、入力症状の正確性が精度に影響することが示唆された。

トップ1の正確率  
 PubCaseFinder : 94% PhenIX : 35%

Licensed under a Creative Commons 表示4.0国際 ©2019 藤原豊史 (DBCLS)

