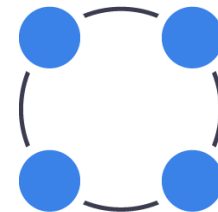


Формирование панели маркеров для молекулярно-генетической диагностики врожденных нарушений обмена веществ

Руководители: Тамара Симакова, Антон Брагин

Студент: Небожатко Екатерина



Цель

Сформировать панель маркеров для секвенирования отдельных генов, ассоциированными с определенными болезнями, и определения степени патогенности мутации.

В панель вошли 37 генов, ассоциированных с 35 различными нозологиями. Для аннотации выбрано 36 внешних источников.

Актуальность

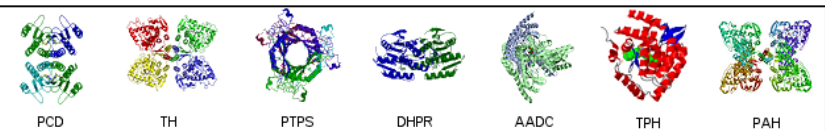
- Предиктивные инструменты, как правило, не обладают 100% чувствительностью
- Для определения степени патогенности мутации можно использовать результаты из различных баз данных, а предиктивные инструменты в качестве дополнения

Задачи

1. Экспортировать данные из внешних библиотек и привести их к единому формату
2. Проаннотировать экспортированные варианты генов инструментами предиктивного анализа рекомендованными ACMG
3. Оценка чувствительности использованных предиктивных инструментов.

Экспорт данных

В базе PNDdb
содержится
информация о генах
GCH1, PTS и QDPR.
База данных
курируется
специалистами



GCH1 Result Details (NG_008647.1)

| | |
|--|--|
| ID No (BIODEF): | |
| ID No (JAKE): | |
| Gene | <i>GCH1</i> |
| Ensembl | ENSG00000131979 |
| Entrez Gene | 2643 |
| HGNC | 4153 |
| Protein Variant | p.GTPCHdel |
| Old name (soment) | |
| Nucleotide aberration cDNA | c.? |
| Nucleotide aberration gDNA | g.108335_132449del |
| Nucleotide Number (use cDNA numbering) | -4000 |
| Codon wild-type | |
| Codon mutant | |
| Sequence Aberration | |
| Variant Type | large del/ins |
| Gene Region | 3'UTR |
| Protein Domain | |
| Enzyme activity | |
| Phenotype | DRD ⓘ |
| Allelic Phenotype (AV) | |
| Splice Taster | |
| Splice Taster Interpretation | not defined |
| Alamut Visual Pathogenicity Report | |
| Experimental_findings | |
| References | Theuns J, Crosiers D, Debaene L, Nuytemans K, Meeus B, Slegers K, Goossens D, Corsmit E, Elinck E, Peeters K, Mattheijssens M, Pickut B, Del-Favero J, Engelborghs S, De Deyn PP, Cras P, Van Broeckhoven C. (2012) Guanosine triphosphate cyclohydrolase 1 promoter deletion causes dopa-responsive dystonia. <i>Mov Disord.</i> 27(11):1451-6. |
| Comments | 24kb deletion spanning exon1 and the 5'UTR |
| Still under investigation | genomic sequence according to NCBI build hg18; AL160471 NO |

<http://www.biopku.org/pnddb/home.asp>

Экспорт данных

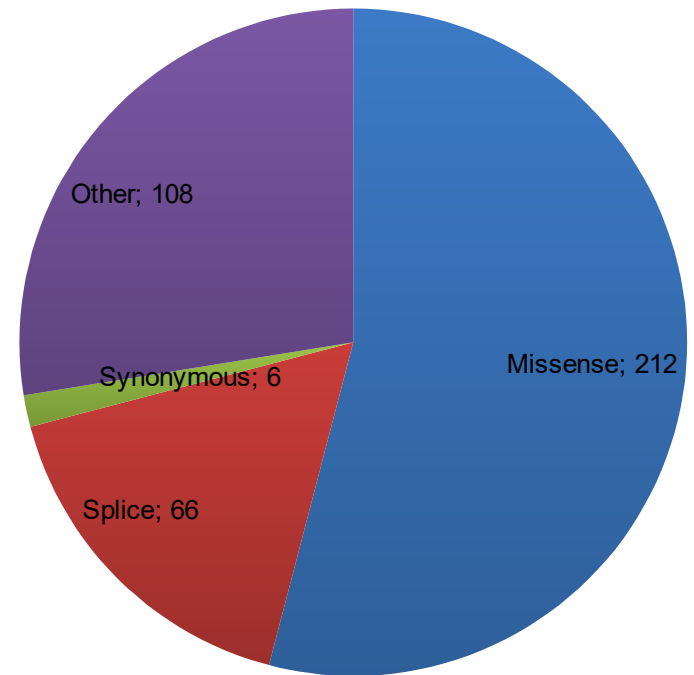
```
##INFO=<ID=ST,Number=0,Type=String,Description="Splice Taster">
##INFO=<ID=SI,Number=0,Type=String,Description="Splice Taster Interpretation">
##INFO=<ID=AV,Number=0,Type=String,Description="Alamut Visual Pathogenicity Report">
##INFO=<ID=EF,Number=0,Type=String,Description="Experimental findings">
##INFO=<ID=RF,Number=0,Type=String,Description="References">
##INFO=<ID=CO,Number=0,Type=String,Description="Comments">
##INFO=<ID=UI,Number=0,Type=String,Description="Still under investigation">
#CHROM POS ID REF ALT QUAL FILTER INFO
chr4 17488779 PNDDB17 G A . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17488816 PNDDB16 C T . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17488828 BIOMDB41 G A . . BI=151|182|255|260|300|390;JA=NA;NG=QDPR;E
chr4 17488835 BIOMDB78 A CCTGTGATCA . . BI=289;JA=NA;NG=QDPR;EN=ENSG00000
chr4 17488839 PNDDB12 GC C . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17488846 BIOMDB371 A G . . BI=706;JA=NA;NG=QDPR;EN=ENSG00000151552;E
chr4 17488854 BIOMDB42 A C . . BI=175;JA=NA;NG=QDPR;EN=ENSG00000151552;E
chr4 17488854 PNDDB32 A G . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17488862 BIOMDB423 TG T . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG
chr4 17492300 BIOMDB43 A C . . BI=145;JA=NA;NG=QDPR;EN=ENSG00000151552;E
chr4 17492304 BIOMDB386 T TT . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG
chr4 17492368 BIOMDB421 CC C . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG
chr4 17492368 BIOMDB53 CC C . . BI=764;JA=NA;NG=QDPR;EN=ENSG00000151552;E
chr4 17493854 PNDDB231 C T . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG
chr4 17493875 BIOMDB420 T TT . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG
chr4 17493885 BIOMDB419 G A . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG
chr4 17493891 BIOMDB418 CC GT . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG
chr4 17493892 BIOMDB63 C T . . BI=193|194;JA=NA;NG=QDPR;EN=ENSG000001515
chr4 17493912 PNDDB19 C T . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17493912 PNDDB85 C G . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17493927 PNDDB20 CC C . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17493927 PNDDB33 T C . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17493928 BIOMDB38 G A . . BI=153|165;JA=NA;NG=QDPR;EN=ENSG000001515
chr4 17493930 PNDDB84 A C . . BI=NA;JA=NA;NG=QDPR;EN=ENSG00000151552;EG=5860;HG
chr4 17493949 BIOMDB47 C T . . BI=198;JA=NA;NG=QDPR;EN=ENSG00000151552;E
```

Код на [github](#)

Был написан скрипт на Python 2.7, реализующий парсинг html страниц и представляющий основную информацию о мутациях в формате VCF.

Аннотация

Для аннотации был использован Variant Effect Predictor. В качестве инструментов предиктивного анализа были использованы SIFT и PolyPhen.



<https://www.ensembl.org/index.html>

SIFT и PolyPhen

PolyPhen2 строит предсказание, основанное на ряде признаков содержащих филогенетическую и структурную информацию о мутации. Использует наивный байесовский классификатор.

SIFT прогнозирует, влияет ли замена аминокислоты на функцию белка на основе гомологии последовательностей и физических свойств аминокислот.

Оценка чувствительности

По нашим данным

- Чувствительность SIFT – 80.9%
- Чувствительность PolyPhen – 73%

По данным статей

- SIFT обладает 83% чувствительностью
([Ngek-Leng Sim et al., 2012](#))
- Polyphen обладает чувствительностью порядка 80% ([Margariga C. Lopes et al., 2012](#))

Выводы

- Реализован инструмент экспорта данных и представлены результаты в VCF формате
- В качестве основного источника определения патогенности мутации лучше использовать информацию из баз данных, а предиктивные инструменты в качестве дополнительного источника аннотации

Спасибо за внимание!