Efficient clinical analysis of DNA sequencing data

Morten C. Eike
Post doc
Dep of Medical Genetics
Oslo University Hospital
Genetic testing

Chromosome

DNA code

Reference

Gene

Variant

Pathogenic
Benign
Uncertain

?
High-throughput DNA sequencing (HTS)

![High-throughput DNA sequencing (HTS)](image)

**Cost of complete genome**

---

**Old tech**

**HTS**

**Moore’s law**

---

**ATGACGGAT**

**AGCCGCA**

**GGATTGGCCGAT**

**TACTGCCATGGCTGCTTGCTTTAACCGCTGTATTT**
Data increase

Old tech

HTS

$1.000 genome - $1.000.000 interpretation
Challenges

• Large amounts of sensitive data
• Complex interpretation process
genAP Norwegian clinical genetic Analysis Platform

- **LAB**: Sequence data
- **EXPERTS**: Manual rules
- **PUBLIC DATA**: In-house DB
- **SEQUENCE DATA**: Raw data
- **ANNOTATED**: Raw data
- **IN-HOUSE DB**: Tjeneste for sensitive data 2.0

- **VM**: Virtual machines

Diagram showing the workflow of genAP, integrating laboratory (LAB) and expert (EXPERTS) processes with public data and an in-house database.
Interpretation
genAP interpreter

- Stepwise process
- Evaluate annotation
- Evaluate references
- Summary and clinical classification
## Rules-based interpretation

### Examples

<table>
<thead>
<tr>
<th>Pathogenic</th>
<th>Benign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein destroyed</td>
<td></td>
</tr>
<tr>
<td>In sick child only</td>
<td></td>
</tr>
<tr>
<td>Small protein change</td>
<td></td>
</tr>
<tr>
<td>Predicted pathogenic</td>
<td></td>
</tr>
<tr>
<td>Predicted benign</td>
<td></td>
</tr>
<tr>
<td>Not inherited</td>
<td></td>
</tr>
<tr>
<td>High population frequency</td>
<td></td>
</tr>
</tbody>
</table>

**Protein destroyed**

- Examples:
  - In sick child only
  - Small protein change
  - Predicted pathogenic
  - Predicted benign
  - Not inherited
  - High population frequency
### -rules from ACMG

<table>
<thead>
<tr>
<th>Classification</th>
<th>Rule</th>
<th>Colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathogenic</td>
<td>(i) 1 Very strong (PVS1) AND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(a) ≥1 Strong (PS1–PS4) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b) ≥2 Moderate (PM1–PM6) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(d) ≥2 Supporting (PP1–PP5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ii) ≥2 Strong (PS1–PS4) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(iii) 1 Strong (PS1–PS4) AND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(a) ≥3 Moderate (PM1–PM6) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b) 2 Moderate (PM1–PM6) AND ≥2 Supporting (PP1–PP5) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(c) 1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5) OR</td>
<td></td>
</tr>
<tr>
<td>Likely pathogenic</td>
<td>(i) 1 Very strong (PVS1) AND 1 moderate (PM1–PM6) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ii) 1 Strong (PS1–PS4) AND 1–2 moderate (PM1–PM6) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(iii) 1 Strong (PS1–PS4) AND ≥2 supporting (PP1–PP5) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(iv) ≥3 Moderate (PM1–PM6) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(v) 2 Moderate (PM1–PM6) AND ≥2 supporting (PP1–PP5) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(vi) 1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5) OR</td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td>(i) 1 Stand-alone (BA1) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ii) ≥2 Strong (BS1–BS4)</td>
<td></td>
</tr>
<tr>
<td>Likely benign</td>
<td>(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ii) ≥2 Supporting (BP1–BP7)</td>
<td></td>
</tr>
<tr>
<td>Uncertain significance</td>
<td>(i) Other criteria shown above are not met OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ii) the criteria for benign and pathogenic are contradictory</td>
<td></td>
</tr>
</tbody>
</table>
# Literature reference evaluation

**Reference evaluation**

Selected reference: Van Hauses et al 2012 (PMID: 23709336)

**Variant:** BREA1 c.387T>C

<table>
<thead>
<tr>
<th>Category</th>
<th>Evaluation</th>
<th>ACMG</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td>Is the reference relevant?</td>
<td>Yes/Indirectly/No/Ignore</td>
<td></td>
</tr>
<tr>
<td>Conclusion</td>
<td>Variant classification</td>
<td>Pathogenic/VUS/Neutral</td>
<td></td>
</tr>
</tbody>
</table>

**Segregation**

- Variant segregates with disease?
  - Strong/Moderate/Weak/No
  - PSX: 3

**Protein**

- Abnormal protein function?
  - ++/+/-/--
  - ...

**RNA**

- Abnormal splicing/protein expression?
  - ++/+/-/--
  - PMX: 2

**In silico**

- Results of prediction tools?
  - Pathogenic/Neutral/Select tool...
  - ...

**Population**

- Increased in affecteds or present in documented healthy individual?
  - RIS>5/Affecteds/Healthy
  - ...

**Gene coverage**

- >90% of gene sequenced?
  - Yes/No
  - ...

**Age of evidence (auto)**

- Reference <10 years?
  - Yes/No
  - ...

**Overall quality**

- Poor/Lacking/Passable/Good/Excellent
  - ...

**Normilised score**

- ...

**Conclusion**

- High quality evidence?
  - Yes/No

**Comments/excerpt of article**

[Comments/excerpt of article]
Rules-based answer

Collect and weigh data:
- annotation
- reference evaluation

Apply weighted rules

Suggested classification with summary
Response from test users:

“When is it available??”
Status

• In place
  – Automated sequence processing and annotation
  – Working prototype of interpreter on TSD 2.0 (wxPython)

• In development
  – Complete rewrite of the interpreter (WebUI+REST API)
  – New rules engine
  – Beta by December 2015

• Test access
  – Medicloud
Some of the people involved

- **Project management:**
  - Thomas Grünfeld (OUS)

- **Steering group:**
  - Dag Undlien (OUS)
  - Margunn Aanestad (IFI)
  - Hans Eide (USIT)
  - Sissel Jor (OUS)

- **Legal issues:**
  - Heidi Bentzen (JUR)
  - Marit Stubø (JUR)

- **Clinical analyses:**
  - Morten C. Eike (OUS)
  - Eidi Nafstad (OUS)

- **System development:**
  - Tony Håndstad (OUS)
  - Espen Skorve (IFI)
  - Svein Tore Seljebotn (OUS)
  - Hugues Fontenelle (OUS)
  - Hallvard Lærum (OUS)
  - Gard Thomassen (USIT)
  - Vegard B. Havdal (Trevis)

**Contact:**
m.c.eike@medisin.uio.no