Disclosure Slide

Financial Disclosure for:
Mari Rossi, MS, PhL
Rare Disease Scientist

Ambry Genetics
Full-time salaried employee

Ambry Genetics
A Konica Minolta Company

ASHG VIRTUAL MEETING 2020 OCTOBER 27-30

Automated HPO terms matching improves but does not eliminate manual review of exome sequencing results

Mari Rossi[1], Kelly Radtke[1], Samin A. Sajan[1], Deepali N. Shinde[1], Jennifer Huang[1], Kelly Hagman[1], Adam Chamberlin[2], Hamed Masnadi-Shirazi[2]

[1]Clinical Diagnostics Department, Ambry Genetics, Aliso Viejo, CA [2]Bioinformatics Department, Ambry Genetics, Aliso Viejo, CA

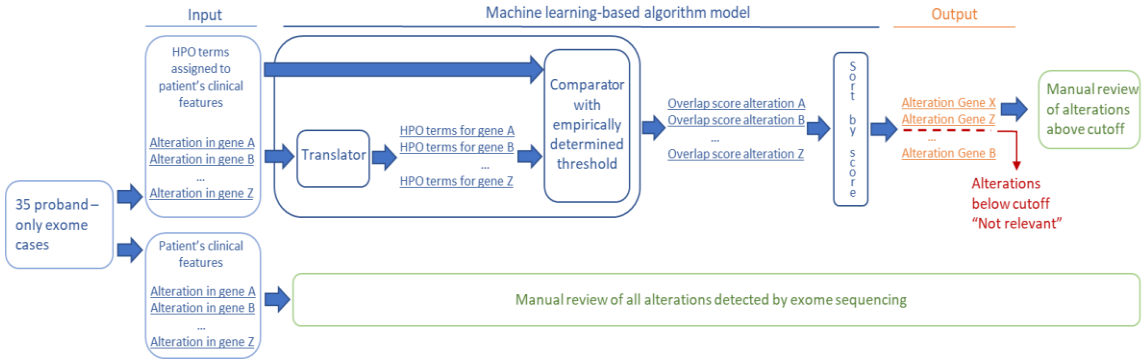
Introduction

Exome analysis is time-intensive and strongly influenced by similarity of patient phenotype with the phenotype associated with a gene. Assessment of clinical overlap can be done either manually or computationally. Utilizing machine learning we optimized a publicly available human phenotype ontology (HPO) database (MGI) to calculate gene phenotypic overlap with the patients' clinical features.

Methods

Manual vs. computational methods for assessing clinical relevance of genes was performed for 35 proband-only cases with exome sequencing at our clinical laboratory (**Figure 1**). Manual time spent on variant review was tracked for each case.

Figure 1: Workflow



Results

Manual review yielded 21 patients with ≥1 reportable alteration in a clinically relevant gene, totaling 28 unique genes.

- Our computational method ranked 24/28 (85.7%) of these genes above the cutoff for clinical relevance, of which 18 (75%) were ranked in the top 5 clinically relevant genes.
- Four relevant genes either ranked below the cutoff or were missed completely by computational methods due to:
  (1) Variable disease spectrum (*MANBA*)
  (2) New -gene-disease association not in the HPO database (*HK1*)
  (3) Ambiguous HPO terms and patient treatment (*GHSR*)
  (4) Lack of HPO terms in the HPO database (*SETD1A*)
- On average, the number of genes manually assessed for each patient was 51 before and 21 after HPO computational ranking.

Table 1: The effects of using computational method on analysis time and false negative results

	Genes reviewed	Relevant genes	Time (min)
Manual review	1738	28	1007
Machine learning-based algorithm model	723	24	714
Improvement	**-58.4 %**	**-**	**-29.1 %**

Table 1: 58.4% of genes were ranked under the cutoff by the algorithm (i.e. clinically irrelevant) which amounted to a 29.1% reduction in total time spent confirming clinical assessment relative to the manual review alone in all 35 patients.

Conclusions

We propose ranking rather than filtering using HPO terms.

- Computational assessment of clinical relevance saves time during exome analysis, but manual review of low-ranking genes is also needed to avoid false negatives.
- Updating the gene-HPO terms database with new gene-disease associations and with separate curations for each gene-disease relationship will minimize these false negatives.

ASHG American Society of Human Genetics