Opening the floodgates to get a sip of water: Challenges of whole exome sequencing analysis as a diagnostic tool


1. Dept. Genetics, UNC-Chapel Hill, 2Dept. Pathology & Laboratory Medicine, UNC-Chapel Hill, 3Renaissance Computing Institute, Chapel Hill, NC

Introduction

North Carolina Clinical Genomic Evaluation by NextGen Exome Sequencing

NCGENES is a research study evaluating whole exome sequencing (WES) as a diagnostic tool in a diverse group of patients with conditions likely to have a genetic etiology, but have evaded diagnosis by traditional methods.

The study aims to answer the following questions:
- Who is the appropriate patient population for (WES)?
- What conditions should be considered for WES analysis?
- Is WES the most efficient & accurate WES analysis method?
- How should incidental or secondary findings be managed?
- Is there an acceptable level of uncertainty in the results for patients/clinicians?

Methods

NCGENES Workflow For Exome Analysis

80K–100K variants/exome

- Gene Filter
- Computational Variant Filters
- Manual Review
- Molecular Sign-out Meeting
- Sanger Confirm & Secondary Review
- Results Returned & Consented to Put Results in EMR

Methods to minimize analysis time

1. A priori Diagnostic Gene Lists: Analysis is limited to broadly designed gene lists consistent with the participant’s disorder.
2. Varies are prioritized by type & effect on protein. Population variants & those with poor quality are filtered out.
3. Manual review of literature, variant databases, allele frequency databases, in silico pathogenicity predictors, etc.
4. Results are discussed weekly by diverse group of clinical geneticists, genetic counselors, clinicians, fellows, etc.

Workflow in CLIA Laboratory

5. Variants meeting our reporting criteria are confirmed via Sanger sequencing.
6. Secondary Variant Review (more thorough): If review alters interpretation, results are discussed at group meeting.
7. Clinical geneticists return results to participants & obtain consent for results to go in EMR (optional).

Participant Demographics

Age of study population

- 29.4%: Children (0-18)
- 33.6%: Adults (19-49)
- 36.8%: Adults (50+)

- Total # of Participants: 575
- Participant age range: 0.64 years
- Participant average age: 32.8 years

Results: First 575 Cases

- Overall Diagnostic Yield:
  - Positive: 22.8% (44)
  - Possible: 18.1% (35)
  - Negative: 59.1% (114)
- Variants requiring analysis, thus limiting VUSs
- Yield varies by age:
  - 18-50yr: 18.0% (38) vs. 21.8% (46)
  - >50yr: 9.5% (16) vs. 21.9% (37)

Diagnostic Yield By Age

- Yield varies by phenotype:
  - IDA &/or Dysmorphology:
    - Negative: 107
    - Possible: 32
    - Positive: 30
  - Ophthalmology:
    - Negative: 86
    - Possible: 30
    - Positive: 15
  - Hereditary Cancer:
    - Negative: 96
    - Possible: 15
    - Positive: 9

- Yield varies by phenotype:
  - IDA &/or Dysmorphology:
    - Highest yield (38.4%) in ophthalmological cases
  - Ophthalmology:
    - Lowest yield (7.6%) in hereditary cancer cases
  - Majority of cases were “Positive” compared to “Possible” except ophthalmological disorders

Acknowledgements

NCGENES is part of the Clinical Sequencing Exploratory Research (CSER) program supported by the National Human Genome Research Institute (NHGRI) & National Cancer Institute (NCI), U31 HG006487 (J.E.F., PI).

References


Conclusions

- How effective is WES as a diagnostic tool?
  - Depends on the condition:
    - Low yield in hereditary cancer
      - Most patients had extensive prior testing for known genes
      - Most often multifactorial, even with a family history
    - Low yield in neuromuscular disorders
      - Conditions in this category tend to overlap with others
      - Often many genes associated with each condition
  - Depends on the age of participant:
    - Lowest yield observed in our cohort over the age of 50
      - Many conditions are more likely to manifest later in life (e.g. cancer, many neuropathies, etc.)
      - Our cancer and neuromuscular cohorts (lowest yields) are mostly comprised of adult participants
      - Verifying variant phase is difficult in elder participants
    - Yield is better in pediatric cohort - where family segregation analysis is most practical

Remaining Challenges

- Variant interpretation is a bottleneck
  - Genome is big & all variation has not been discovered
  - Large majority of variants will be VUSs
  - Rare variants are frequent & difficult to assess
  - Use of appropriate filters can help reduce the number of variants requiring analysis, thus limiting VUSs

- Limited phenotypic information in the clinic
  - Directly impacts interpretation of results
    - Difficult to narrow the list of variants with limited clinical information
    - Difficult to differentiate between diagnostic & incidental findings
  - Propose that clinical labs work closely with clinicians

- Can we successfully balance benefit vs. harm?