

Establishing an approach to reduce Sanger confirmation in NGS-based studies

P48



Nuno Teixeira^{1,2}, Joana Moreira¹, Alexandra Lopes¹, Mariana Ferreira¹ and Rita Cerqueira¹

¹CGC Genetics/Unilabs, Molecular Diagnostics and Clinical Genomics, Porto; ²FCUP- Faculty of Sciences of the University of Porto, Porto

BACKGROUND

Clinical diagnostic testing has been revolutionized by Next-Generation Sequencing (NGS). This powerful technology has been replacing Sanger sequencing, the gold standard methodology for clinical molecular diagnosis¹. Currently, and given the complexity of NGS, no specific guidelines are addressing NGS quality parameters or providing concrete guidance for confirmatory analysis of NGS results². Sanger sequencing is used as a broad methodology to validate each reported variant, however, these confirmations on all NGS-detected variants not only increase the test cost, but also delay the turnaround time³. Some studies¹⁻⁴ propose that independent confirmations of NGS findings are unnecessarily redundant and should not be required for high-quality variants. In this context, the aim of this work is to establish internal quality thresholds, above which Sanger confirmation of NGS variants will not be necessary.

METHODOLOGY

A retrospective analysis of 500 Single Nucleotide Variants (SNVs), detected by NGS-based studies and further confirmed by Sanger sequencing, was conducted to determine quality thresholds based on specific variant metrics. Libraries were prepared using an oligonucleotide-based target capture chemistry by Agilent Technologies. Illumina next-generation sequencing was performed, and DRAGEN Enrichment was used for alignment and base calling.

500 SNVs



VCF -Variant Call Format

Variant features:

Type of variant;
Location;
Read Depth;
Call Quality;
Allele frequency;
Zygosity;
Position;
Allelic depth;
Mapping quality;
...

RESULTS

The 500 SNVs under study were obtained from 459 samples patients and are found in 289 different genes.

The data used to establish the quality thresholds are shown in figures 1 and 2. A decision tree was built considering these variant-associated features (Figure 3).

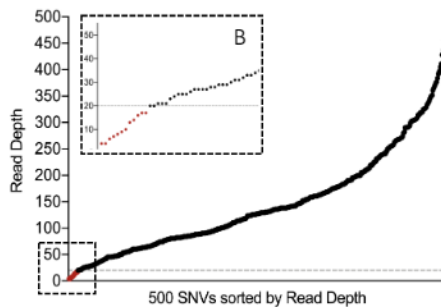


Figure 1 - Results sorted by Reading Depth (RD). Each SNV is represented by a point, sorted by ascending quality score. Red points represent SNVs with RD < 20 (horizontal grey dashed line), representing variants that require Sanger confirmation. B) ZOOM of the cut off region, highlighted the 12 SNVs with RD < 20.

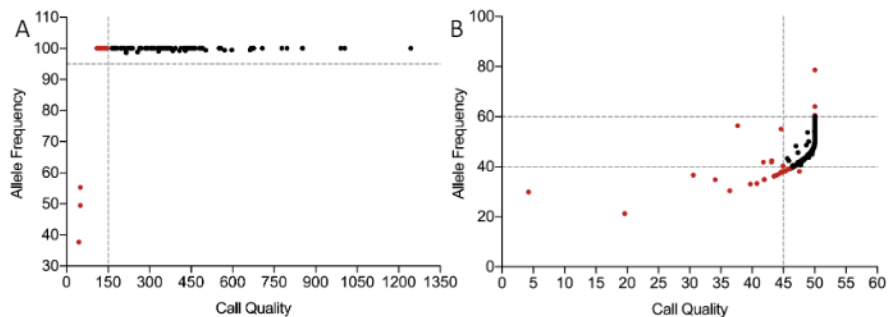


Figure 2 - The Allelic Frequency (AF) and Call Quality (CQ) profile of the 488 SNVs with RD ≥ 20. A- For Homo-/Hemizygous variants, a dashed gray vertical line marks the CQ threshold of 150, while the AF threshold is 95, represented by the horizontal dashed gray line. B- For heterozygous variants, a lower threshold, equal to 40, and an upper threshold equal to 60 are shown as horizontal gray dashed lines. The QC threshold is 45 and is shown as a grey dashed vertical line. The red points represent variants that do not meet at least one of the criteria. For Homo-/Hemizygous variants: AF < 95 or QC < 150; for Heterozygous variants: AF is out of the interval 40 to 60 or QC < 45. The red points in both graphs represent variants that require Sanger confirmation.

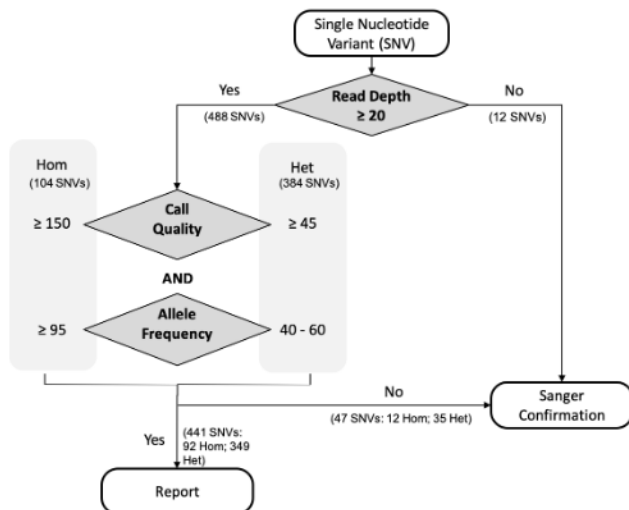


Figure 3 - Figure 3 - Workflow for high-quality variants identification. A decision tree was built to identify high-quality SNVs based on specific variant metrics. The number of filtered SNVs, assessed in this study is cited in brackets.

- Of the 500 total SNVs analyzed, 488 variants had a Read Depth (RD) ≥ 20, representing 98% of the total variants (Figure 1). From these 488 variants:
 - Homozygous variants - Call quality ≥ 150 and an AF ≥ 95. Of the 104 homozygous variants, there are 92 meet these criteria (Figure 2A).
 - Heterozygous variants- Call quality ≥ 45 and an AF between 40 and 60. Of the 384 heterozygous variants, 349 variants meet these criteria (Figure 2B).

CONCLUSION

- Only 59 SNVs do not meet these filtering criteria and, therefore, should be confirmed by Sanger sequencing.
- Sanger sequencing confirmed 99.6% of the NGS variants.
- Based on this study, Sanger confirmation of SNVs will be reduced by approximately 89% (Figure 3).
- Although Sanger confirmation of low-quality SNVs and all other variants remains necessary, this new approach will lead to a significant decrease in costs and labor.

REFERENCES

- ¹ Strom, S. P. et al. Assessing the necessity of confirmatory testing for exome-sequencing results in a clinical molecular diagnostic laboratory. *Genet. Med.* **16**, 510-515 (2014);
- ² Baudhuin, L. M. et al. Confirming Variants in Next-Generation Sequencing Panel Testing by Sanger Sequencing. *J. Mol. Diagn.* **17**, 456-461 (2015);
- ³ Mu, W., Lu, H. M., Chen, J., Li, S. & Elliott, A. M. Sanger Confirmation Is Required to Achieve Optimal Sensitivity and Specificity in Next-Generation Sequencing Panel Testing. *J. Mol. Diagn.* **18**, 923-932 (2016);
- ⁴ Bauer, P. et al. Development of an evidence-based algorithm that optimizes sensitivity and specificity in ES-based diagnostics of a clinically heterogeneous patient population. *Genet. Med.* **21**, 53-61 (2019).