



# An automatic end-to-end solution for disease-causing variant detection in rare and hereditary diseases with a high case solve rate and a much reduced false positive rate

A. Joecker<sup>1</sup>, S. Shah<sup>2</sup>, M. Hadjisavas<sup>2</sup>, D. Richards<sup>2</sup>, B. Oester<sup>1</sup>, G. Eley<sup>3</sup>, B. Solomon<sup>3</sup>, J.G. Vockley<sup>3</sup>

<sup>1</sup>QIAGEN Aarhus, Aarhus, Denmark; <sup>2</sup>QIAGEN Silicon Valley, Redwood City, USA; <sup>3</sup>Inova Translational Medicine Institute, VA, USA

## Motivation & Goals

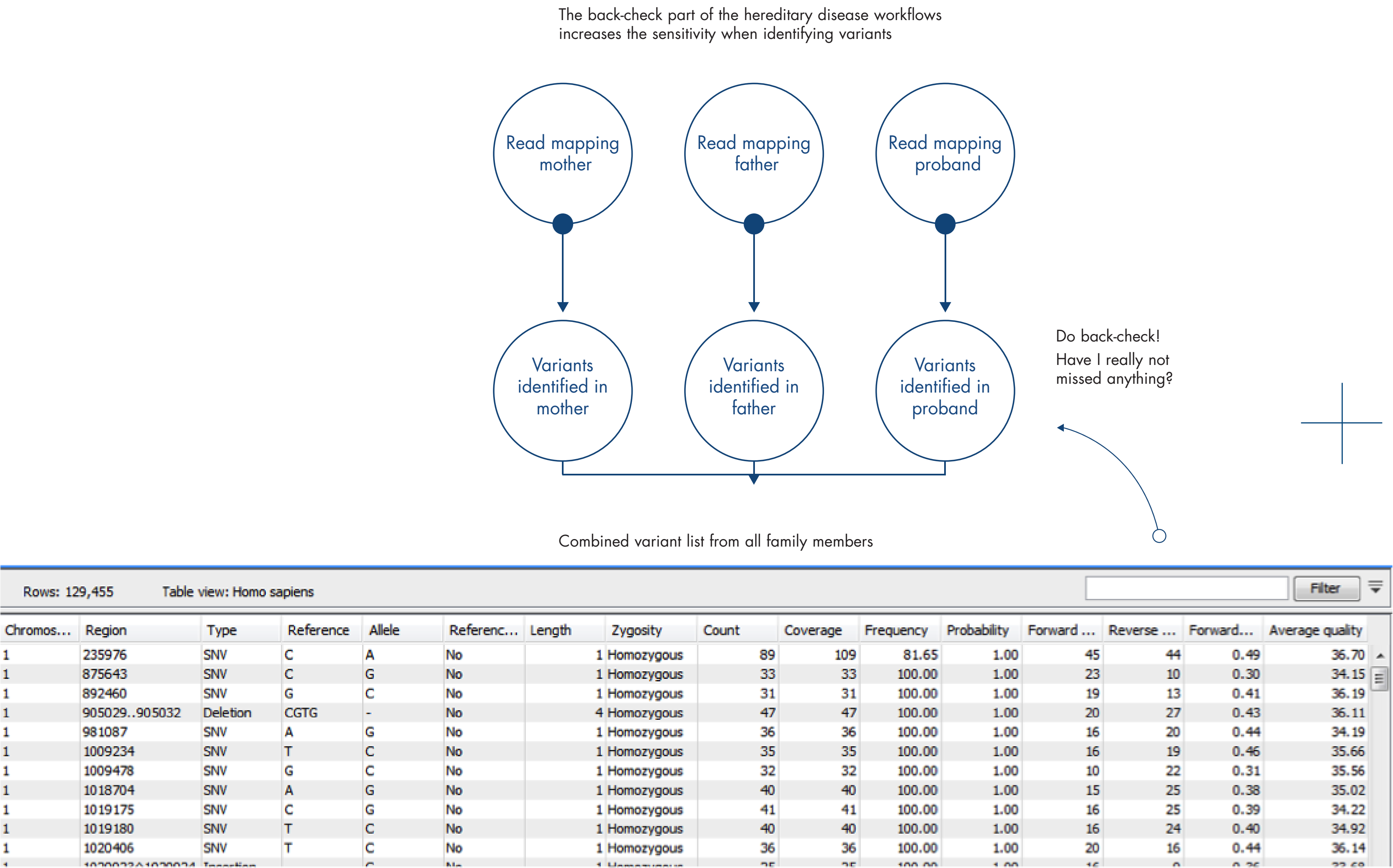
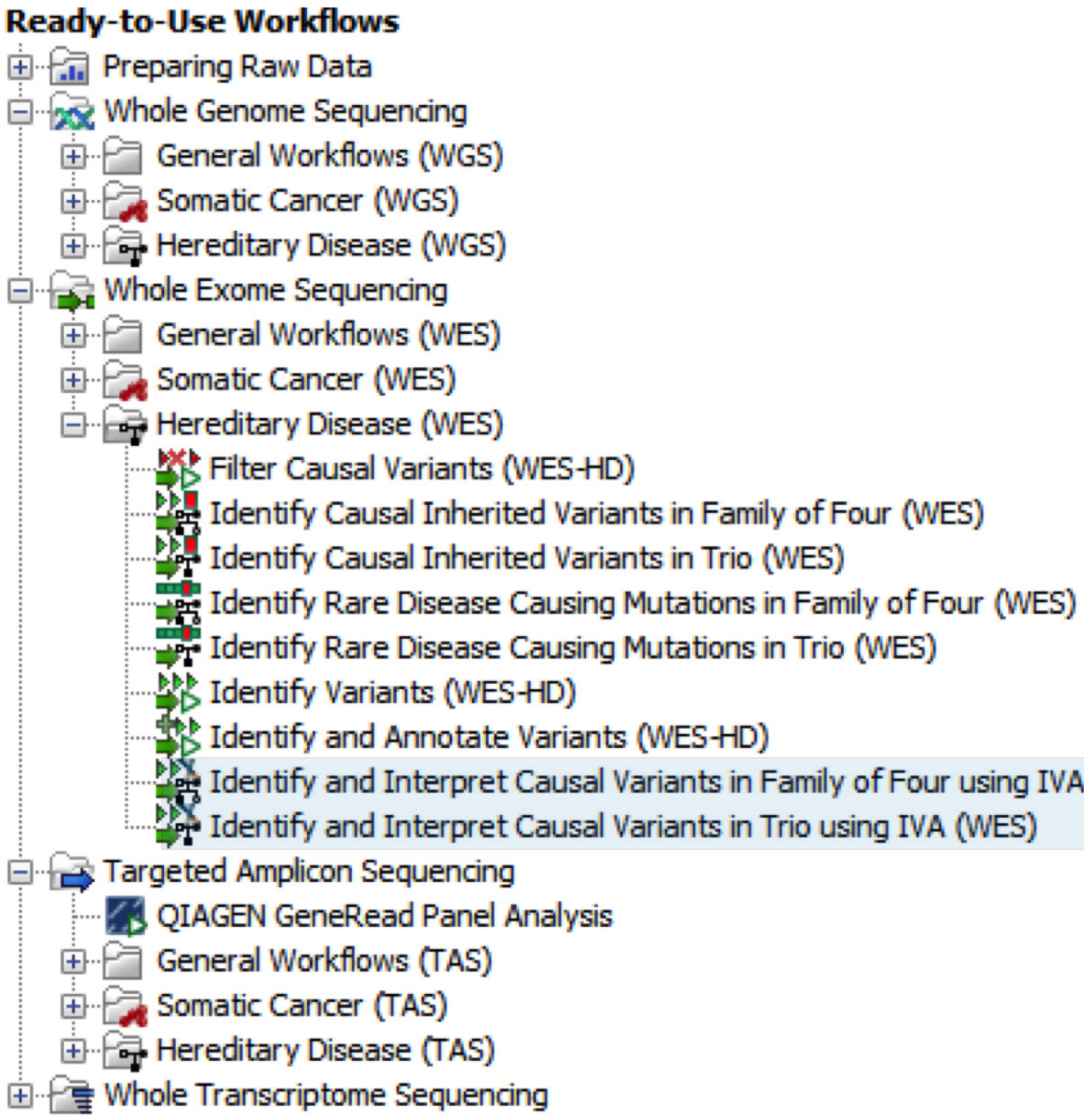
Identification of causal variants in hereditary diseases can be both challenging and time consuming. Often a lot of time and resources are wasted on identifying a disease causing variant from a TRIO and trying to validate variants, which are actually not disease causing or artifacts. Furthermore, sometimes no causal variant can be identified at all in the patient.

QIAGEN Bioinformatics' hereditary disease solution delivers increased sensitivity for identifying causal variants, while reducing significantly the list of candidate variants for follow-up. In addition, it is very easy to run as data analysis and interpretation steps are embedded in a streamlined end-to-end workflow with optimized parameter settings.

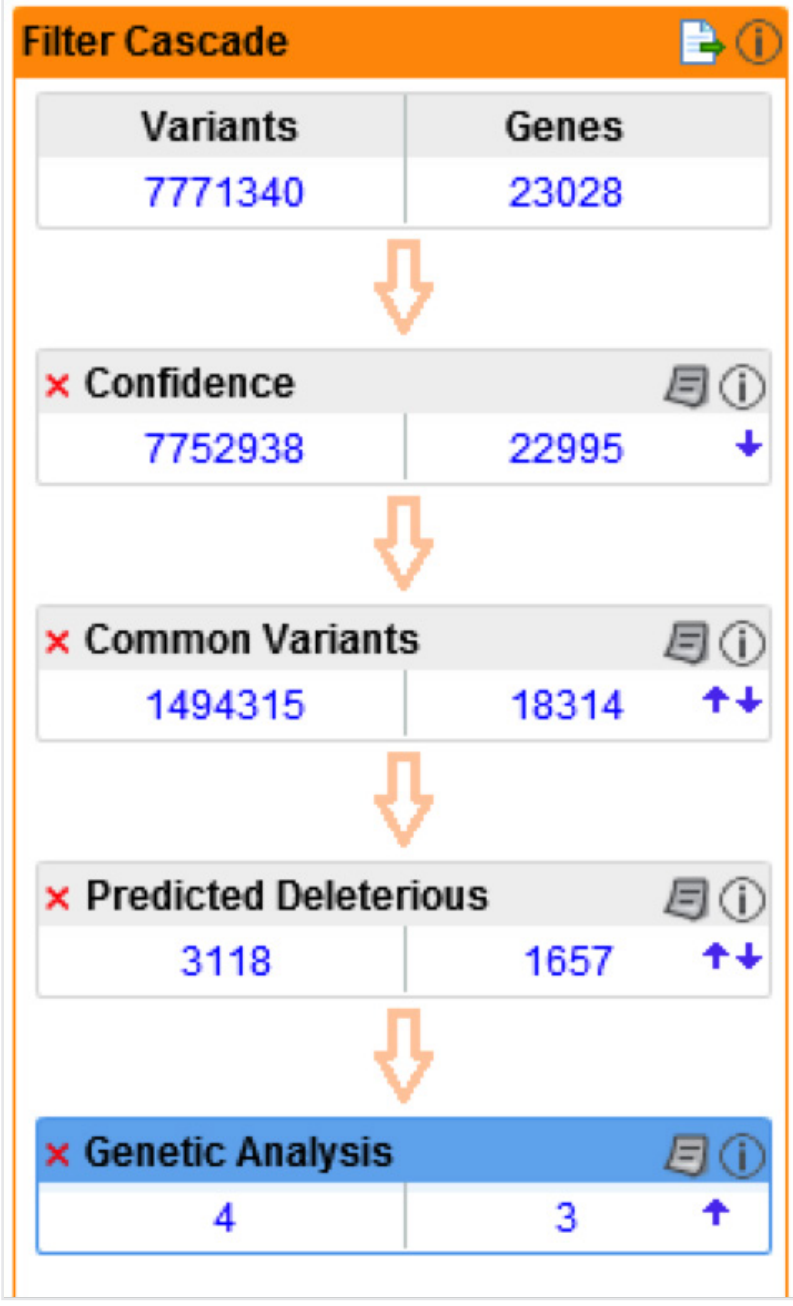
In this study we show the first results from a benchmarking study with six whole genome trios and one whole exome trio.

## Materials & Methods

Hereditary Disease end-to-end workflows in Biomedical Genomics Workbench [BxWB]

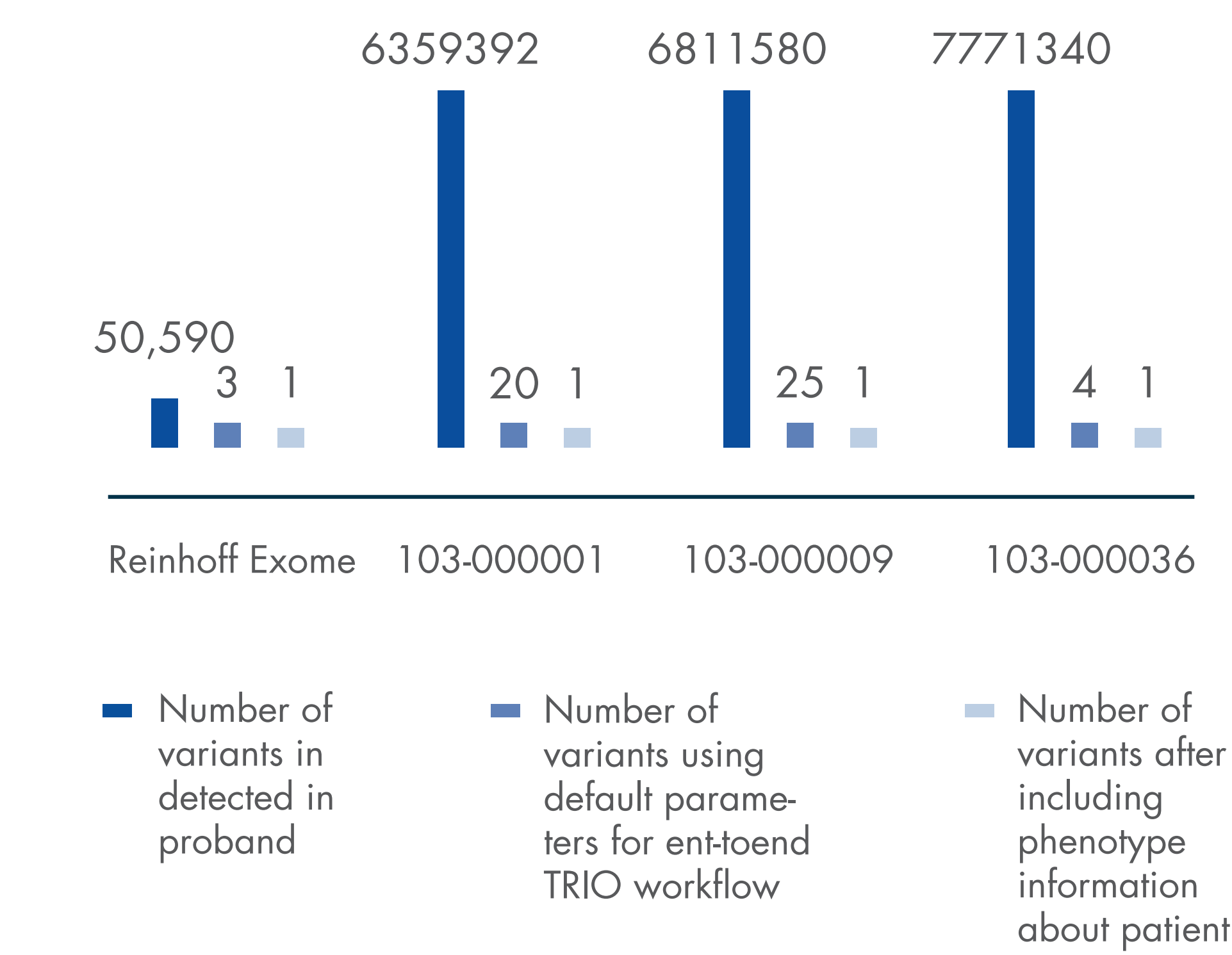


Filter cascade in Ingenuity Variant Analysis



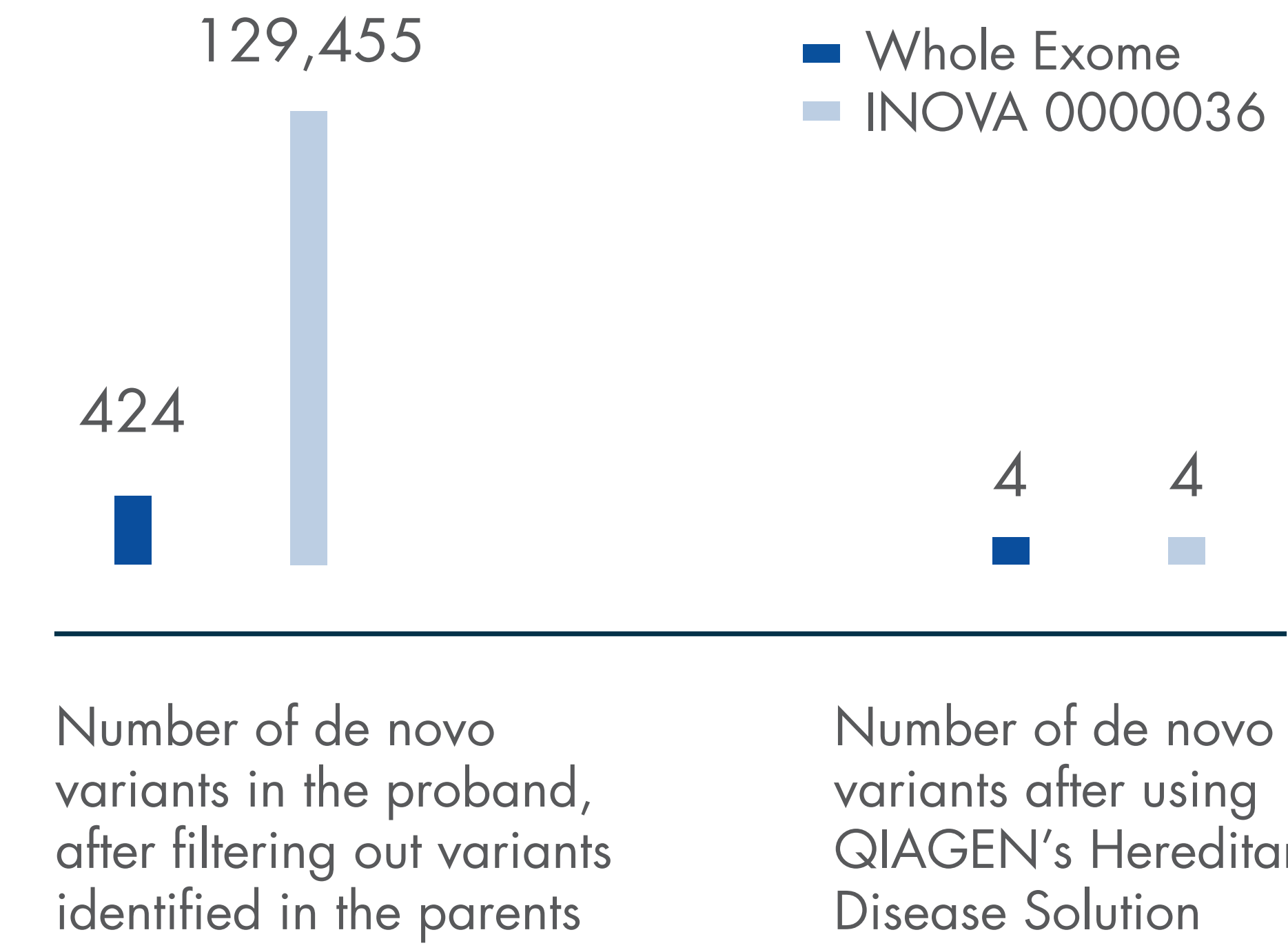
These two workflows are part of the Ingenuity Variant Analysis plugin available for Biomedical Genomics Workbench and Biomedical Genomics Server Solution. They perform an end-to-end analysis-to-interpretation workflow and were used for the study here on six whole genome trios of the INOVA Genomes and one whole exome datasets<sup>1</sup> for which the disease causing variants are previously known.

## QIAGEN's hereditary disease solution for TRIOs is able to shorten the list of candidate variants to one candidate in case of a *de novo* inheritance pattern



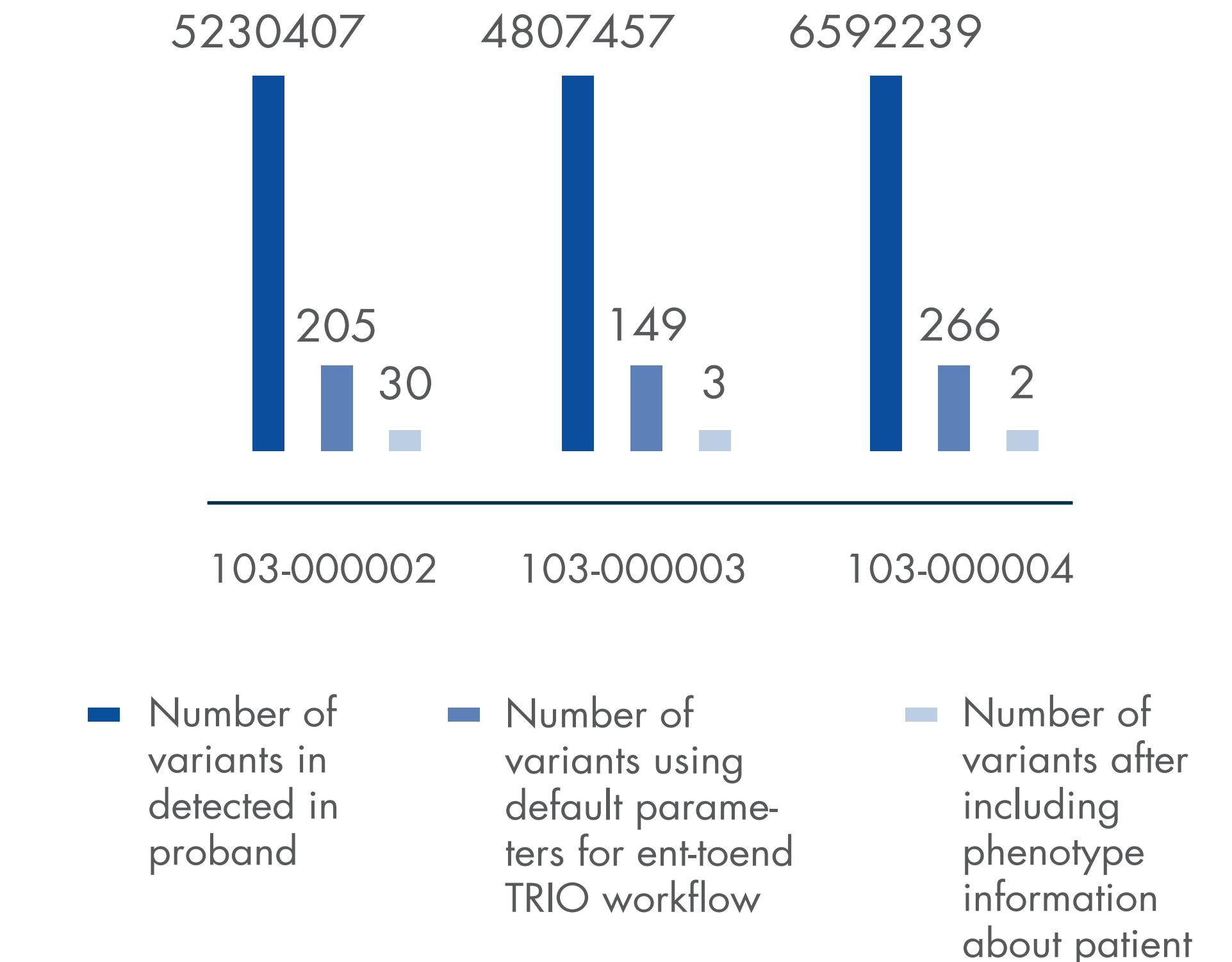
In all cases the previously identified disease causing variant was in the end result. In three of the whole genome trios and the exome dataset, the cause of the disease was a *de novo* mutation. In these cases the disease causing variant could be reported as the only candidate variant in the result when phenotype information from the physician during examination of the child where included in the filter cascade in Ingenuity Variant Analysis. The complete workflow was run with default parameters.

## The number of false positive *de novo* variants in the child can be reduced by 99%



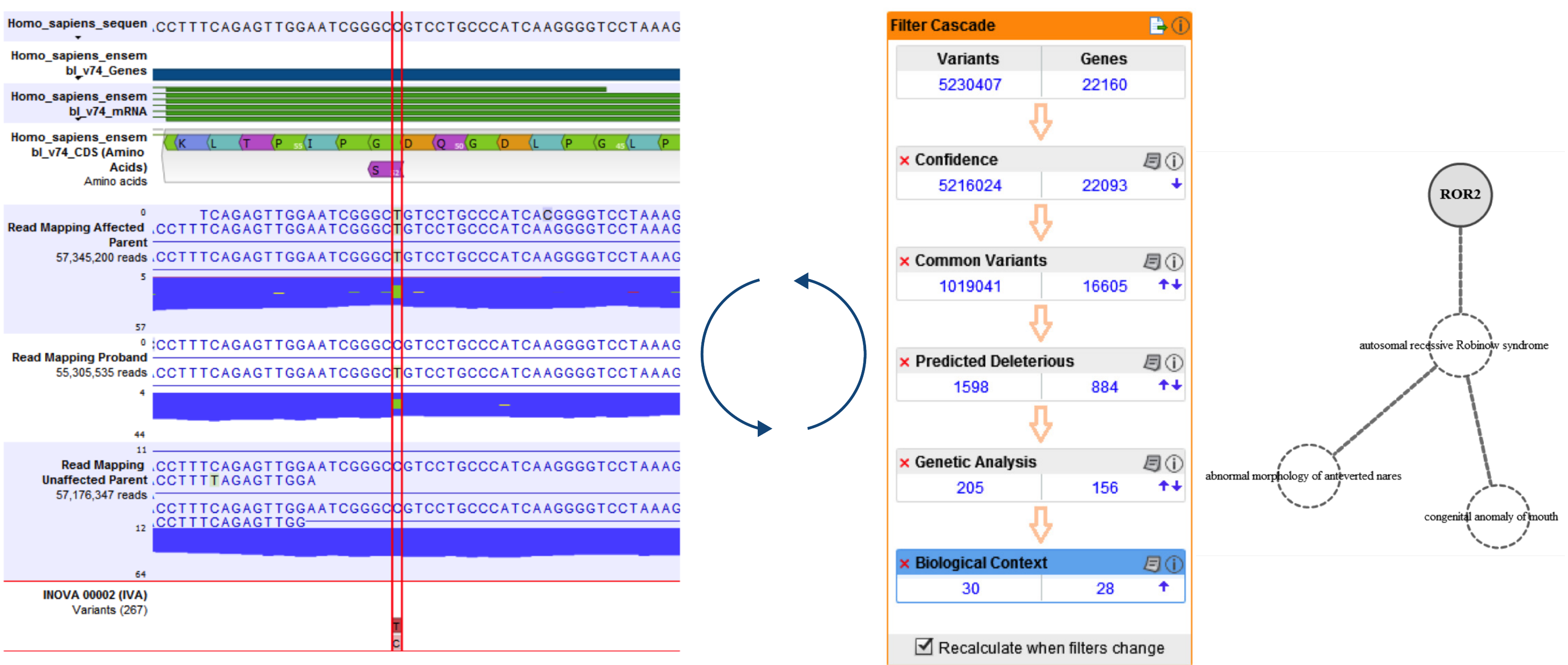
A common approach to the discovery of *de novo* variants is based on the variant call results of all family members. To identify *de novo* variants that potentially are disease causing, all variants present in mother and father are subtracted from the list of variants found in the child. Due to low read coverage and allelic dropout, many of the variants are usually not called in all individuals, which leads to a high number of false positives. The workflow with Biomedical Genomics Workbench, Biomedical Genomics Server Solution and Ingenuity Variant Analysis removes this bias in the investigated whole exome and whole genome trio.

## QIAGEN's workflow reduces the list of candidate variants by 99% in case of dominant inheritance



Also in case of dominantly inherited variants was QIAGEN's Hereditary Disease solution able to detect the disease causing variant candidates. Moreover, it was able to shorten the list of candidate variants for follow-up by more than 99% when phenotype information was provided. In two whole genome trios the list of candidates could even be shorten to less than 5 candidates. In one case the final result included 30 candidates for follow-up, of which two were potentially be disease causing.

## An easy-to-use, but still flexible solution



The result is achieved by seamlessly integrating Biomedical Genomics Workbench, Biomedical Genomics Server Solution and Ingenuity Variant Analysis in a one step workflow. Results are shown in a Genome Browser View, enabling easy validation of a variant. Right clicking on the variant track opens the filter cascade in Ingenuity Variant Analysis.

It can be optimized to include for example phenotype information in the analysis by adding a Biological Context Filter. Updated results can afterwards be fetched from Ingenuity Variant Analysis. They then can again be visualized in the Genome Browser View in Biomedical Genomics Workbench. The example above shows the results of the analysis of the 103-000002 INOVA whole genome trio. Among the 30 candidate causal variants found by the workflow, and taking into account the disease phenotype, ROR2 is the most promising one.

## Summary and Discussion

In this study we analyzed six whole genome and one whole exome TRIO using QIAGEN's new hereditary disease solution. The end-to-end workflow in Biomedical Genomics Workbench includes a checking back into mapped sequencing reads and optimal filter settings for Ingenuity Variant Analysis. This results in a high case solve rate and shortens the list of disease causing candidate variants to a minimum. We show that for diseases caused by *de novo* variants, the complete workflow results in the identification of the disease causing variant only without any additional candidates. This is achieved using default parameters and providing phenotype information to the filter cascade. On dominant inherited diseases we were able to reduce the number of candidates for follow-up by more than 99%. In addition, we are allowing the easy validation of the candidates and an optimization of the filter cascade. As a result less time and resources have to be spent on additional validation steps.

## References

1. Rienhoff HY Jr. et al. (2013) A mutation in TGF $\beta$ 3 associated with a syndrome of low muscle mass, growth retardation, distal arthrogryposis and clinical features overlapping with Marfan and Loeys-Dietz syndrome. Am J Med Genet A. 2013 Aug;161A(8):2040-6