VIRTUAL GENETICIST™

Full Suite of AI-Powered Tools for Rare Disease Diagnosis

Platform Overview

Virtual Geneticist™ is an advanced AI-powered bioinformatic platform designed specifically for the diagnosis of rare genetic disease. The platform works with both whole genome and whole exome data and features a range of AI-powered tools that drive up diagnostic yields while drastically reducing the time required for variant interpretation. In head-to-head comparisons and in studies by clinicians at British Columbia's Children's Hospital, Virtual Geneticist™ (VG) was able to accurately rank causal variants in minutes and outperform competing platforms in the market[1]. To achieve this goal, VG deploys proprietary AI technologies to provide insights for accurate variant classification including the intelligent scanning and weighing of relevant scientific publications and the detailed evaluation of patient phenotypes and genomic sequencing data.

Virtual Geneticist™ has been validated for clinical use and has been integrated into standard NGS workflows in a number of CAP and CLIA-certified genomics laboratories as well as pediatric medical centers and research facilities. Built with maxi-parallel architecture, the platform can process dozens of cases simultaneously and enables a range of applications from the hyper-efficient re-analysis of previous cases to deployment within large scale population-level genomic research and pharmaceutical discovery.

Solves the Bottleneck of Clinical Variant Interpretation

In recent years, breakthroughs in technology have drastically reduced the costs and availability of both whole exome and whole genome sequencing. Unfortunately, the process of data interpretation remains a significant barrier to the widespread adoption of these tests. A recent study reports that even with commercial bioinformatic software, a single case may require a highly-qualified, primarily PhD-level staff member to spend nearly 8 hours to complete the analysis for just one patient test[2]. As such, the process of clinical interpretation represents by far the mostly costly and time-consuming aspect in clinical genetic testing. Adding to the dilemma is the fact that average diagnostic yields for clinical whole genome and whole exome tests remain below 50%[3], leaving far too many cases unsolved.

Despite these challenges, the accurate diagnosis of rare genetic disease is a healthcare priority both in the U.S. and around the world. It is estimated there are roughly 6000 known rare genetic diseases that effect over 250 million people globally, the majority of whom are children. Even in the U.S. and the U.K. where the healthcare systems are relatively well-funded, a recent survey found that rare disease patients typically visited 8 physicians, received between two and three misdiagnoses, and endured an average of 6 years before receiving the correct genetic diagnosis[3].

Best-in-Class Performance for Ranking of Disease-causing Genetic Variants

Accurately Ranks Disease-causing Genetic Variants

Figure 1. Retrospective analysis of rare disease pediatric cases utilizing singleton WES samples. Study led by Dr. Adrienne Elbert, MD - British Columbia Children's Hospital (n = 219). Results presented at the 2023 American Society of Human Genetic Conference (ASHG).

BREAKTHROUGH GENOMICS

VIRTUAL GENETICIST™

Accurate Genomic Data Interpretation in Minutes and at Scale

Outperforms Other Software Platforms in the Market

In 2023, VG was at the center of a retrospective study conducted by clinicians at BC Children's Hospital to establish best practices on variant interpretation within their specialty clinical group which sees some of Canada's most challenging genetic disease cases. In their study involving 219 previously reported clinical WES cases performed by U.S. commerical labs, VG outperformed both the Exomizer and also the Lyrical platform in its ability to accurately rank and prioritize the correct causal variant within a Top 10 List (Figure 2). In addition, Virtual Geneticist presented the correct variant in the #1 ranked position 79% of time with both the Exomizer and Lyrical performing below 55% (Figure 3).

Figure 2

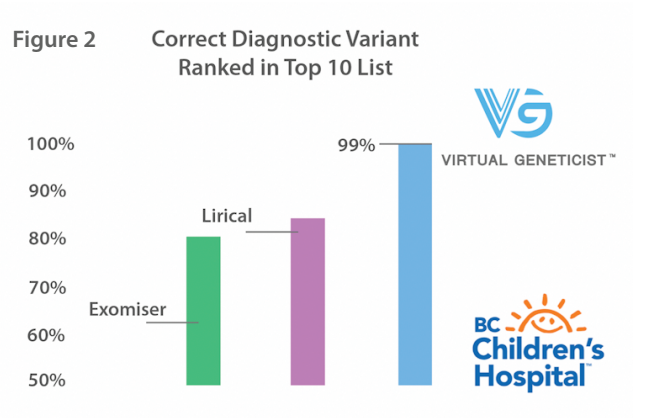
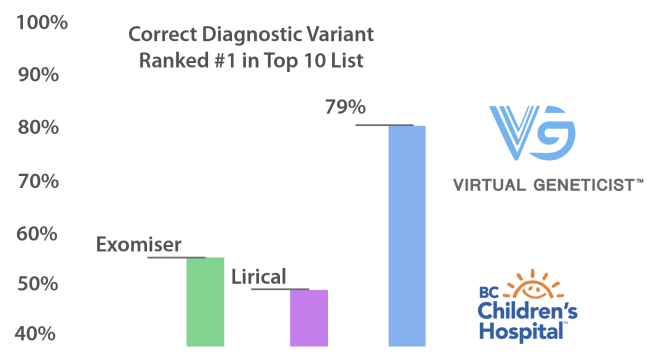


Figure 3

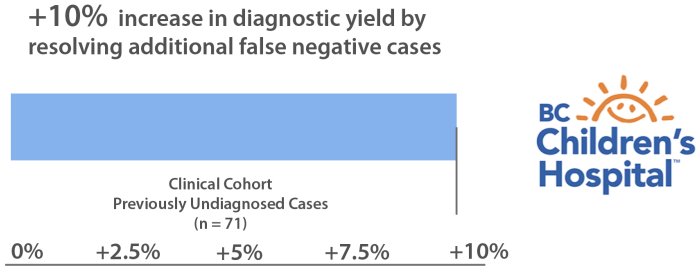


Comparative Study of Variant Randing and Prioritization Results from 3 bioinfomatic platforms utilizing singleton WES rare disease pediatric cases (n=219)[1].

Immediate Re-Analysis Increases Diagnostic Yield by Resolving Previously Reported False Negative Cases

In another part of the study, 71 previously undiagnosed clinical cases that had been returned to the hospital by commercial labs without a diagnosis were then re-analyzed utilizing the VG platform. The team found seven new diagnoses for pediatric caes that had been missed for an increased diagnostic yield of 10% (Figure 4). This significant result has clear real-world implications, especially for the individual patients and their families who rely on an accurate diagnosis for decisions regarding disease management and treatment.

Figure 4



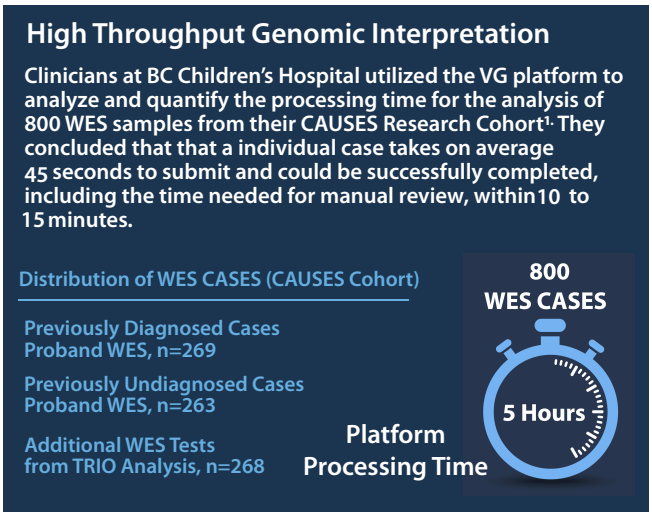
In her discussion on the results, Dr. Adrienne Elbert, MD who led the study, concluded that VG was able to successfully take into account blended phenotypes and dual diagnoses leading to many of the new findings in these complex cases.

Another aspect of the BC Children's Hospital Study was that the team was able to run 800 WES samples through the platform within just 5 hours of total of processing time. To achieve true high-throughput tertiary analysis, VG leverages advanced cloud-based architecture that enables multi-parallel processing combined with the integration of high-speed GPUs.

BREAKTHROUGH GENOMICS

VIRTUAL GENETICIST™

Figure 5



AI-Driven Variant Prioritization and the Resolution of Variants of Unknown Significance (VUS)

VG's advanced variant ranking is accomplished through proprietary machine-learning based algorythms that take into account both genomic and clinical information and weigh dozens of critical variables. The overall process is set up specifically to follow the classification guidelines set out by the American College of Medical Genetics (ACMG). Within the ranking of each variant, VG calculates 3 unique numerical scores. These scores have values from 0 to 1 and correspond to the 3 categories of analysis within the platform (Phenotype Score - P Score, Genome Score - G Score, and Final Score - which takes into account all factors). These proprietary scores help users evaluate the relative strength of evidence in the classification and prioritization of the different variants. Also unique to Virtual Geneticist is the platform's fully-interpreted Variant Literature Database which intelligently extracts critical information from each paper and includes those elements in the calculations for each variant.

World's First Fully-Interpreted Variant Literature Database with Over 10 Million Variants Containing Concise and Structured Summaries

An Intelligent and Detailed Literature Overview for Every Variant

An essential part of variant curation is the review of relevant publications for each variant being considered. To address this previously unmet need, Virtual Geneticist includes a one-of-a-kind database that helps users quickly assess the most critical information from published papers. The information contained in each concise literature entry includes gene-disease associations, patient-variant information, experimental and functional studies, and family-based segregation analyses that are all required for variant classification based on the guidelines set out by the American College of Medical Genetics (ACMG). The database utilizes a proprietary AI-driven genetic language model that has resulted in comprehensive and verified data extracted from over 11 million variant-paper pairs.

Accurate and Fully-Automated ACMG Classifications for over 180 Million Variants

The platform's powerful AI-based engine has classified over 180 Million variants according to ACMG Guidelines. Accurate classifications with easy-to-read supporting evidence presentations (Figure 6) are calculated from both publicly available data and detailed variant information extracted from each variant-paper pair.

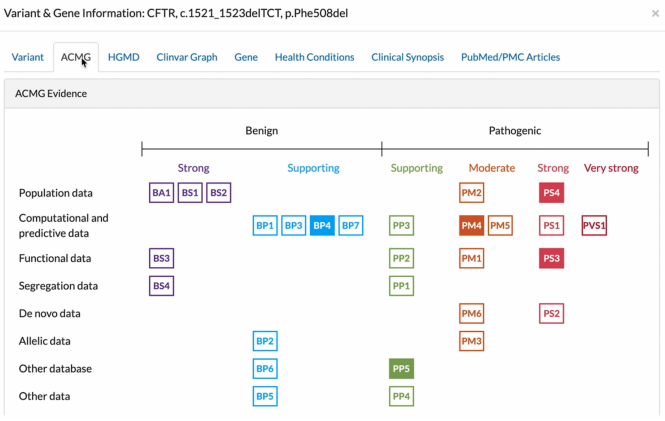


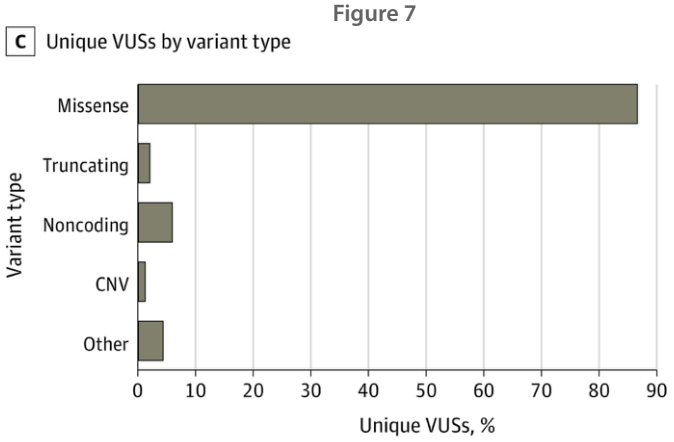
Figure 6. Virtual Geneticist's ACMG Classification Evidence Presentation makes is easy for users to see the relative strength of evidence calculated for each variant classification.

VIRTUAL GENETICIST™

Predicts Missense and other Categories of Challenging Genetic Variants

Recent large-scale studies emphasize the increasingly prevalence and reporting challenges associated with VUS variants[11]. Variants in this category include missense variants (by far the most common), truncating variants, noncoding variants, copy and structural number variants (CNVs) and others that often require the use of specialized bio-informatic tools and techniques to better understand their potential impact on patients[11].

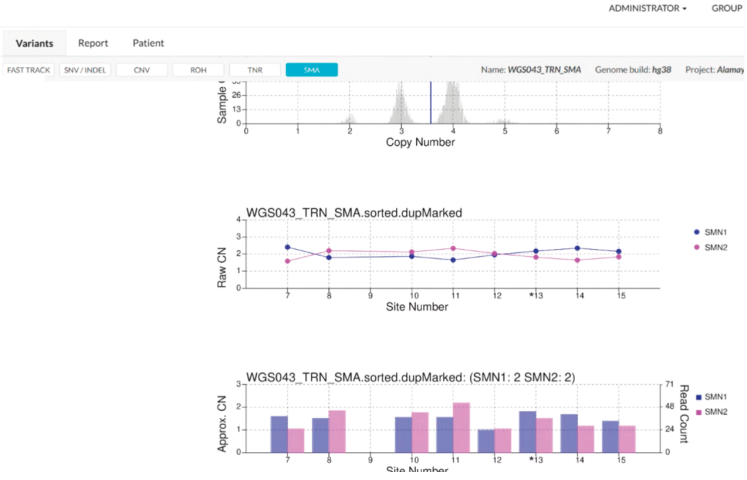
Figure 7



Chen E et al, 2023, PMID: 37878314

Provides Advanced Tools to Analyze CNVs and Structural Variants, Trinucleotide Repeats, UPDs, and More

Additional capabilites provided by VG include analysis and prioritization of structural variants and cnvs, short tandem repeats, trinucleotide repeat expansions, and the ability to identify previously difficult-to-diagnose genetic conditions such spinal muscular atrophy (See Figure 8 below for Virtual Geneticist's SMA Detection Module)



Streamlined User Interface Without Complex Onbarding or Instruction

Many other bioinformatic platforms currently available for genomic interpretation are often characterized by their users as having "steep learning curves" and requiring "complicated onboarding procedures".

In contrast, VG was designed with a simple user-based approach with additional time-saving tools for both the input and selection of clinical notes and patient phenotypes, as well as a customizable and self-populating reporting module to make the generation of clinical reports more efficient while reducing the potential for manual errors.

The VG platform is sequencer agnostic and accepts both whole exome and whole genome data from a variety of sequencers including Illumina, MGI, and others. VG is also unique in that tools for both secondary and tertiary analysis are integrated into one platform so there is no need for additional hardware or expensive bio-informatic plug-ins.

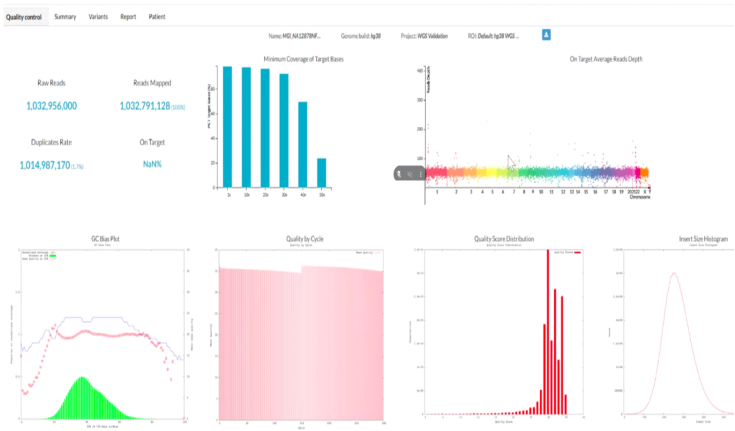


Figure 9. Virtual Geneticist's Dashboard for Secondary Analysis and Quality Control (QC). The platform enables the user to set customizable alerts with specific thresholds that can be set to help monitor and flag potential deficiencies within the underlying NGS sequencing data.

BREAKTHROUGH GENOMICS

VIRTUAL GENETICIST™

Powers Re-Analysis of Past Cases and Cases with Previously Negative Findings

A. Elbert 20231104

Enables the Efficient Re-Analysis of Previous WES and WGS Cases

A number of studies have demonstrated that periodic re-analysis of rare disease cases can also have a significant impact diagnostic yields[2]. Re-analysis can be especially effective with updates to the patient's phenotype information or as new research and functional studies are published on specific variants[2] allowing clinician's to update their findings as needed.

This was also one of the central conclusions of the study conducted by BC Children's Hospital. Namely, that internal and immediate re-analysis of unsolved cases using VG did significantly improve their diagnostic yields without significant investment in time or additional infrastructure. Furthermore, Dr Elbert reported that using VG was, in essence, multimodal since it allowed the user to explore unreported variants while at the same time gain an understanding of the potential impacts that adjusting the phenotype terms can have on the suggested diagnosis and variant ranking[1]. She also stressed how the re-analysis conducted within the study did not depend on the evolution of knowledge on a given variant or the change in a patient's phenotype over time, but was based entirely on the increased sensitivity and overall performance of the Virtual Geneticist Platform.

At the end of her talk at ASHG, Dr Elbert further emphasized how important it is in clinical practice to have alternative methodologies like Virtual Geneticist available to clinicians to double-check the results returned by commercial labs and alternative pipelines.[1]

Questions answered:

1) **Can internal re-analysis of unsolved cases improve diagnostic rates?**
   YES! False negative rate of ~10%
2) **Can an AI-based tool (Virtual Geneticist) help to identify missed diagnoses?**
   YES!
   Virtual Geneticist is powerful in detecting disease-causing variants, even in complex cases.
   With Virtual Geneticist, the clinician can:
   - Explore unreported variants/ new diagnoses
   - Explore phenotype/ variant ranking relationships
   - Contextualize reported VUS
   - Contextualize non-diagnostic results
3) **What factors lead to missed diagnoses?**
   Differences in integration of data about the variant, phenotype, family, disease, and its mechanism, and thresholds for reporting

Figure 10. Final Slide with overall conclusions from Dr. Elbert's Presentation at the 2023 American Society of Human Genetics Conference (ASHG)

Supports Breakthroughs in Clinical Genomics and Novel Variant Discovery

Deployed in various research settings, VG has also been listed as an important tool in a number of recent scientific studies in the field of medical genetics. The studies that credit Virtual Geneticist include: the discovery of new causal variants in congenital heart disease[6], alternative tandem splice mechanisms in children with developmental delay[10], and a groundbreaking functional study of PACS1 variants linked to syndrommic intellectual disability disorder[9]. Two additional papers that also utlize VG focus on the guiding principles behind variant interpretation[8] and the weighing of evidence to support reported genomic variation in clinical practice[7]. (Figure 11)

Figure 11

Recent Publications Crediting Virtual Geneticist for Assiting in New Findings

Genetics in Medicine

AMERICAN JOURNAL OF medical genetics

frontiers | Frontiers in Medicine

BREAKTHROUGH GENOMICS

References

BREAKTHROUGH GENOMICS

1. Improving Transparency, Collaborative Variant Interpretation, and High-throughput Diagnostics: A Clinician's Review of the VIRTUAL GENETICIST Platform. Study results and talk presented at the 2023 American Society of Human Genetics Conference by Dr. Adrienne Elbert, MD. British Columbia Children's Hospital

2. Austin-Tse, C.A., Jobanputra, V., Perry, D.L. et al. Best practices for the interpretation and reporting of clinical whole genome sequencing. npj Genom. Med. 7, 27 (2022). https://-doi.org/10.1038/s41525-022-00295-z

3. Chung CCY; Hong Kong Genome Project; Chu ATW, Chung BHY. Rare disease emerging as a global public health priority. Front Public Health. 2022 Oct 18;10:1028545. doi: 10.3389/f-pubh.2022.1028545. PMID: 36339196; PMCID: PMC9632971.

4. Bagger, F.O., Borgwardt, L., Jespersen, A.S. et al. Whole genome sequencing in clinical practice. BMC Med Genomics 17, 39 (2024). https://doi.org/10.1186/s12920-024-01795-w

5. Marwaha, S., Knowles, J.W. & Ashley, E.A. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. Genome Med 14, 23 (2022). https:doi.org/10.1186/s13073-022-01026-w

6. Di Francesco D, Swenerton A, Li WL, Dunham C, Hendson G, Boerkoel CF. Are CUL3 variants an underreported cause of congenital heart disease? American Journal of Medical Genetics. September 2023. https://doi.org/10.1002/ajmg.a.63387

7. Chin HL, Gazzaz N, Huynh S, Handra I, Warnock L, Moller-Hansen A, Boerkoel P, Jacobsen JOB, du Souich C, Zhang N, Shefchek K, Prentice LM, Washington N, Haendel M, Armstrong L, Clarke L, Li WL, Smedley D, Robinson PN, Boerkoel CF. The Clinical Variant Analysis Tool: Analyzing the evidence supporting reported genomic variation in clinical practice. Genezt Med. 2022 Jul;24(7):1512-1522. doi: 10.1016/j.gim.2022.03.013. Epub 2022 Apr 19. PMID: 35442193; PMCID: PMC9363005

8. Handra J , Elbert A , Gazzaz N , Moller-Hansen A , Hyunh S , Lee HK , Boerkoel P , Alderman E , Anderson E , Clarke L , Hamilton S , Hamman R, Hughes S , Ip S , Langlois , Lee M , Li WL , Mackenzie F , Patel M. , Prentice L. , Sangha K , Sato L, S , Seppelt M , Swenerton A, Warnock L , Zambonin J. , Boerkoel CF. , Chin H , Armstrong L. The practice of genomic medicine: A delineation of the process and its governing principles. Frontiers in Medicine. 2023 Vol.9, https://doi.org/10.3389/fmed.2022.1071348

9. Moller-Hansen A, Duha H, Hyun KL, Jenea B, Yunhan Y, Kun C, Li WL, Thomas G, Boerkoel, CF. Do PACS1 variants impeding adaptor protein binding predispose to syndromic intellectual disability? American Journal of Medical Genetics. 2023, May. https://doi.org/10.1002/ajmg.a.63232

10. Sage A, Hyun K, Dalmann J, Lin S, Samra S, Salman A, Del Bel K, Li WL , Lehman A, Turvey A, Boerkoel CF, Richmond P. Generation of tandem alternative splice acceptor sites and CLTC haploinsufficiency: A cause of CLTC-related disorder. American Journal of Medical Genetics. 2023 May. https://-doi.org/10.1002/ajmg.a.63249

11. Chen E, Facio FM, Aradhya KW, Rojahn S, Hatchell KE, Aguilar S, Ouyang K, Saitta S, Hanson-Kwan AK, Capurro NN, Takamine E, Jamuar SS, McKnight D, Johnson B, Aradhya S. Rates and Classification of Variants of Uncertain Significance in Hereditary Disease Genetic Testing. JAMA Netw Open. 2023 Oct 2;6(10):e2339571. doi: 10.1001/jamanetworkopen.2023.39571. PMID: 37878314; PMCID: PMC10600581.