SURVEY

AI-driven software platforms for genomic analysis and interpretation

Artificial intelligence (AI) has assumed an important role in clinical and genomic diagnostics (D'Agaro 2018; Dias and Torkamani 2019; Sennaar 2019), and a recent report from the PHG Foundation examines "*the drivers behind the recent rise of AI techniques for genomics, existing and emerging applications, the limitations of AI for genomic medicine, and the challenges to realising its full potential for health*" (Raza 2020).

A number of companies offer AI-driven software platforms for genomic analysis and interpretation of clinical sequencing data (e.g., NGS, WES, WGS), using, for example, VCF files as input (Table 1). Analysis tasks include alignment, variant interpretation, variant calling, annotation and analysis, and literature curation. Advantages of the AI-driven approach include greatly reduced turn-around-time and increased diagnostic yields. There are also AI-based variant calling algorithms (some freely-available) based on supervised learning (e.g., ISOWN), machine learning (e.g., BAYSIC, MutationSeq, SNooPer, SomaticSeq), convolutional neural networks (e.g., Clairvoyante), deep convolutional neural network (e.g., DeepSea), deep recurrent neural network (e.g., Deep Nano), deep neural network (e.g., DANN), and artificial neural networks (e.g., Skyhawk), and these have been surveyed and critiqued recently (Bohannan and Mitrofanova 2019; Karimnezhad et al 2020; Koboldt 2020; Liu et al 2019; Xu 2018).

Table 1. AI-driven software platforms for genomic analysis.

Company and website	Description
Binartis https://www.binartis.com	BINOME© platform for AI-based genetic variant interpretation.
Breakthrough Genomics https://btgenomics.com	Proprietary machine learning algorithm, Enlighter™ provides an end-to-end genomic analysis platform for whole exome, whole genome, and gene panel tests. Ranks and prioritizes variants, real-time literature presentation for each variant, free-text phenotype filter, copy number and structural variant analysis.
Broad Institute https://gatk.broadinstitute.org/hc/en-us	GATK, machine-learning, genome analysis toolkit focused on variant discovery [e.g., identify germline copy number variants, somatic short variants (SNVs and Indels), germline short variants (SNPs and Indels)].

Diploid http://www.diploid.com/moon	Software package (Moon) that autonomously diagnoses rare diseases from NGS data using AI.
emedgene https://www.emedgene.com	Cognitive Genomic Intelligence™ machine learning algorithms provide a shortlist of causative variant, curates evidence for every variant identified, and provides links to the supporting literature databases
enGenome https://www.engenome.com	eVai combines AI with ACMG, AMP and ClinGen guidelines, and classifies and prioritizes every genomic variant for pathogenicity, suggesting all the possible related genetic diagnosis.
Fabric Genomics (formerly Omicia) https://fabricgenomics.com	Comprehensive platform for NGS analysis, interpretation and clinical reporting. Identifies and prioritizes disease causing variants in rare idiopathic disease, enables clinical reporting for hereditary disease panels, accelerates WGS interpretation to diagnose genetic disorders.
Geneyx https://geneyx.com	Clinical genomics platform for WGS providing management, analysis and interpretation (e.g., phenotype-driven variant prioritization, automated ACMG/ClinGen variant classification) of genetic data, and clinical reporting.
Genomenon https://www.genomenon.com/mastermind	Mastermind, AI-driven genomics search engine for variant interpretation. Identifies every genomic association in the medical evidence, drawing informative connections between genes, variants, diseases, phenotypes, therapies, copy number variations, and categorical keywords to inform clinical care**.**
Genoox https://www.genoox.com	A cloud-based advanced AI framework (encompasses purpose-built applications) for managing the entire genetic sequencing process and delivery of clinically actionable insights and disease diagnoses.
Genuity Science https://genuitysci.com	Applies domain-specific AI algorithms to reveal novel patterns and causal dependencies. Provides interpretation and actionable insights for whole genome

	sequencing, whole exome sequencing, or transcriptome sequencing.
Genosity https://www.genosity.com (acquired by Invitae: https://www.medicaldevice-network.com/news/invitae-to-acquire-genosity/)	Case Analyzer variant review and reporting platform. The AI-based Genosity engine built in to learn as the data grows to facilitate faster analysis.
Google https://github.com/google/deepvariant	DeepVariant, an open-source convolutional neural network for variant calling from next-generation DNA sequencing data.
Lifebit https://lifebit.ai	Clinical/multiomics data analysis with Lifebit AI Engine (part of Lifebit CloudOS Marketplace advanced interpretation tools).
Nostos Genomics https://www.nostos-genomics.com	Variant interpretation platform for automated variant classification based on proprietary machine learning.
Ocean Genomics https://oceangenomics.com	AI-driven platform (txome.ai) extracts high dimensional transcriptomic features from RNA-seq data and combines it with clinical metadata to learn models and identify candidate multidimensional biomarkers.
Omicia - see Fabric Genomics	-
OmniTier https://www.omnitier.com	CompStor Novos® genomics analysis appliance for assembly-based variant calling using a tiered-memory algorithm. CompStor Insight™ NGS tertiary analysis employing machine learning-based scoring.
Sentieon https://www.sentieon.com	DNAscope for germline SNV/INDEL Variant Calling using machine learning enhanced filtering for top variant calling accuracy. TNscope for somatic SNV/INDEL Variant Calling using machine learning enhanced filtering.
SIVOTEC BioInformatics http://www.sivotecanalytics.com	GENA, a GENome Analysis tool, part of the AI-driven, integrated SIVOGEN platform of tools for clinical interpretation of SNP array results.
SOPHiA https://www.sophiagenetics.com	SOPHiA for Genomics is a collective artificial intelligence that analyzes complex NGS data by detecting,

	annotating, and pre-classifying SNVs, Indels, and CNV (SOPHiA DDM® platform). Provides clinical-grade genomic solutions for accurate detection and characterization of genomic variants associated with cancers and hereditary disorders. SOPHiA continuously learns from thousands of patients' genomic profiles and experts' knowledge to improve patients' diagnostics and treatments.
Variantyx https://www.variantyx.com/technology/	Genomic Intelligence® platform algorithms harnesses AI and big data to uniquely identify and analyze all major types of variants in WGS output.

References

Bohannan ZS, Mitrofanova A. Calling variants in the clinic: Informed variant calling decisions based on biological, cinical, and laboratory variables. Comput Struct Biotechnol J 2019;17:561-9. https://doi.org/10.1016/j.csbj.2019.04.002.

D'Agaro E. Artificial intelligence used in genome analysis studies. EuroBiotech J 2018;2:78-88.

Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med 2019;11:70. https://doi.org/10.1186/s13073-019-0689-8.

Karimnezhad A, Palidwor GA. Thavorn K *et al.* Accuracy and reproducibility of somatic point mutation calling in clinical-type targeted sequencing data. BMC Med Genomics 2020;13:156. https://doi.org/10.1186/s12920-020-00803-z.

Koboldt DC. Best practices for variant calling in clinical sequencing. Genome Med 2020;12:91. https://doi.org/10.1186/s13073-020-00791-w.

Liu Z, Zhu L, Roberts R, Tong W. Toward clinical implementation of next-generation sequencing-based genetic testing in rare diseases: Where are we? Trends Gen 2019;35:852-67. https://doi.org/10.1016/j.tig.2019.08.006.

Raza S. Artificial intelligence for genomic medicine. PHG Foundation. University of Cambridge. 2020. https://www.phgfoundation.org/documents/artifical-intelligence-for-genomic-medicine.pdf.

Sennaar K. Machine learning in genomics – Current efforts and future applications. Bus Intell Anal November 28, 2019. https://emerj.com/ai-sector-overviews/machine-learning-in-genomics-applications/.

Xu C. A review of somatic single nucleotide variant calling algorithms for next-generation sequencing data. 2018;16:15-24. https://doi:org/10.1016/j.csbj.2018.01.003.