VarElect – Powerful gene-based NGS biological interpretation for revealing potential variants playing a role in disease

Following Next Generation Sequencing (NGS), base calling, sequence mapping to the reference genome and variant calling are performed (Primary and Secondary analysis), and a Variant Call Format (VCF) file is generated. For a whole exome sequence this file typically contains ~50,000 variant rows. Tertiary analysis is then applied with the global aim of identifying a few variants underlying a disease/phenotype of the sequenced subject or cohort.

In a first stage of Tertiary analysis, annotation is performed, whereby the VCF file is populated with new columns with descriptors and scores for every variant. These include variant type and frequency in general populations, the gene within (or near) which the variant occurs, and the predicted functional impact of missense variants on the encoded protein.

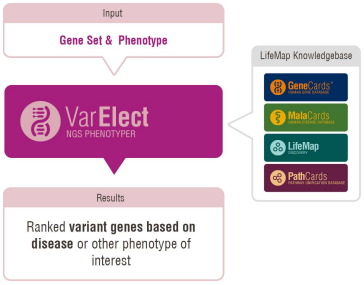
Sequential or concomitant filtering via the annotation parameters is then performed, leading to a list of a few dozen to a few hundred variants, depending on the assumed mode of inheritance and on the employed filtering cutoffs, such as frequency in controls and predicted protein impact. A typical result could be 300 variants with population frequency below 0.3%, all with “damaging” protein impact score.

At this stage the goal is to identify disease/phenotype relationships for the variants. In a minority of the cases this can be attained via further annotation by public disease variant databases such as ClinVar and COSMIC. But since a typical gene can have thousands of variants that have not yet been documented to have a disease relationship, there is a low probability that such lookup will successfully find the coveted relations between any of the variants and the specific disease/phenotypes of the sequenced subject.

At this point a gene-based interpretation process becomes necessary. The LifeMap strategy entails seeking a disease/phenotype relationship for the *gene* that harbors the VCF variant rather than for the variant itself.

VarElect, LifeMap Sciences’ NGS biological interpretation tool, thus offers a powerful gene-based interpretation process to reveal potential variants playing a role in disease diagnosis and/or treatment. VarElect is powered by LifeMap’s GeneCards Suite knowledgebase, a robust searchable set of integrated databases

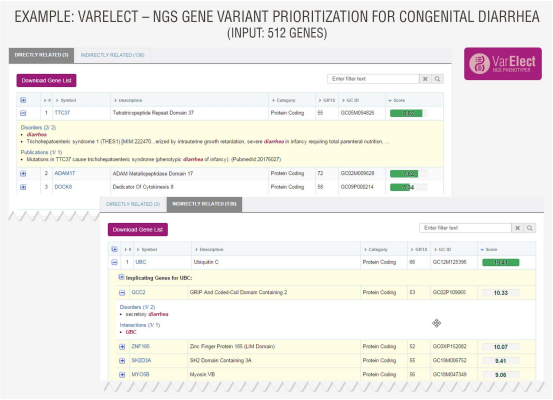
capable of revealing relations among any of tens of thousands of genes to any of >100,000 terms that signify diseases names, aliases, phenotypes and symptoms.



In LifeMap's VarElect-powered biological interpretation process, the filtered or complete list of genes which harbor the variants is subjected to an efficient search, capable of scoring the existence and strength of relation(s) between the gene and the supplied disease/phenotype/symptom terms.

The result is a ranked list of genes, the top of which contains, with high likelihood, the best disease gene candidates. Because every gene in the list is related, via the VCF file, to the original variant discovered, VarElect's interpretation process results in informed, unambiguous relationships between all sequence-discovered variants and concrete disease descriptors of the subject.

In the screen shot on the right, an example of VarElect results is shown. Variant-harboring genes related to the disease/phenotype are ordered by association scores, along with supporting evidence.



VarElect can also be used in a VarAnnot™ mode, whereby the aforementioned scores become additional annotation columns in the VCF file. As such, they can be used in a powerful, one-shot joint filtration with other previously obtained annotation entries.

In summary, the goal in NGS is identifying variants related to a subject's disease/phenotype. VarElect addresses a central challenge in this process by effectively performing phenotype prioritization for documented and undocumented variants. As such, it emerges as an indispensable tool in NGS biological interpretation.

Interested in trying VarElect? Sign up for a free trial at VarElect.GeneCards.Org

VarElect
NGS PHENOTYPER