

POSTER PRESENTATION

Open Access

# Exome sequencing combined with semantic discovery identifies strong disease-associated candidates in a single case of relapsing remitting multiple sclerosis

Mahjoubeh Jalali Sefid Dashti<sup>1</sup>, Maritha Kotze<sup>2</sup>, Susan Janse van Rensburg<sup>2</sup>, Alan Christoffels<sup>1</sup>, Junaid Gamielidien<sup>1\*</sup>

From Beyond the Genome 2012

Boston, MA, USA. 27-29 September 2012

## Background

As known disease-associated variants identified through large cohort-based studies often explain only a small percentage of genetic risk in multifactorial disorders such as multiple sclerosis (MS), alternative methods for identification and prioritization of variants that directly and/or indirectly play a role in disease development have become increasingly important. We were tasked with identifying possible genetic causes in a case of atypical relapsing remitting MS (RRMS) that also presented with porphyria-like symptoms and where demyelination was halted in the patient upon iron supplementation. As the patient had no parents or siblings that could be used as references for filtering exome variants, we aimed to develop a new prioritization strategy based on the combination of a predicted deleterious effect on the protein and existing knowledge of the biological roles of the genes and their contribution to relevant phenotypes.

## Materials and methods

Exome sequencing was performed and functional SNP analysis and frameshift indel detection were carried out using a combination of prediction tools. Predicted deleterious variants were further assessed with respect to their possible involvement in MS using our internal biomedical semantic database, the B.O.R.G. (BioOntological Relationship Graph), which integrates existing biomedical knowledge and uses path-based graph theoretic querying to discover links between biological concepts. The semantic model used in this study incorporated known human,

mouse and rat: gene, gene-to-disease, gene-to-phenotype, gene-to-function and gene-to-pathway and orthology relationships in conjunction with 'surrogate' phenotype and biological function links to the relevant disease term.

## Results

The exome sequence resulted in identification of 64,890 variants based on the human reference genome, of which 4,847 missense variants were predicted to have a damaging effect. These were used to interrogate the semantic network, which simplified exploring the network for transitive relationships that may explain the biological contribution of identified mutations to the development of disease, either directly from human evidence or transitively via model organism evidence such as knockout phenotypes. 750 variants were found to be potentially involved in MS based on known gene-to-disease associations or via surrogate-phenotype, -function or -pathway links to MS. Examples of strong candidate MS genes identified were: *BACE1*, *HEXB* and *NRCAM* (involved in myelination), *CNTN2* (involved in axon development), and *CCM2* (implicated in immune response and central nervous system inflammation). In addition, we identified deleterious variants in *IREBP* and *CYBRDI*, which are involved in iron regulation and homeostasis, and may contribute to the iron deficiency condition of the patient. Several other strong candidates are being evaluated further.

## Conclusions

While the list of candidates obtained from this study is quite large, a number of them may in fact represent a part of the large number of variants proposed to be associated with the 'missing' (approximately 80%) MS genetic risk,

<sup>1</sup>South African National Bioinformatics Institute, University of the Western Cape, Bellville, 7535, South Africa  
Full list of author information is available at the end of the article

with each adding a tiny percentage to the overall risk of developing the disease. As these variants fulfil many criteria, our knowledge-driven prioritization strategy may appear to have the potential to improve discovery of causative variants in non-Mendelian diseases and also rare diseases where large cohorts are almost impossible to build.

#### Author details

<sup>1</sup>South African National Bioinformatics Institute, University of the Western Cape, Bellville, 7535, South Africa. <sup>2</sup>Department of Pathology, University of Stellenbosch, Tygerberg, 7505, South Africa.

Published: 1 October 2012

doi:10.1186/1753-6561-6-S6-P10

**Cite this article as:** Jalali Sefid Dashti *et al.*: Exome sequencing combined with semantic discovery identifies strong disease-associated candidates in a single case of relapsing remitting multiple sclerosis. *BMC Proceedings* 2012 **6**(Suppl 6):P10.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

