

ORAL PRESENTATION

Open Access

is-rSNP: a novel technique for in silico regulatory SNP detection

Geoff Macintyre^{1,2*}, James Bailey^{1,2}, Izhak Haviv^{3,4,5}, Adam Kowalczyk²

From Sixth International Society for Computational Biology (ISCB) Student Council Symposium
Boston, MA, USA. 9 July 2010

Determining the functional impact of non-coding disease associated SNPs identified by genome-wide association studies (GWAS) is challenging. Many of these SNPs are likely to be regulatory SNPs (rSNPs): variations which affect the ability of a transcription factor (TF) to bind to DNA. However, experimental procedures for identifying rSNPs are expensive and labour intensive. Therefore, in silico methods are required for rSNP prediction. By scoring two alleles with a TF position weight-matrix (PWM), and observing the change in PWM score, it can be determined which SNPs are likely rSNPs. Predicting in this manner, however, yields large numbers of false positive predictions. In addition, no method exists that determines the statistical significance of a nucleotide variation on a PWM score.

We have designed an algorithm for in silico regulatory SNP detection called is-rSNP [1]. We employ novel convolution methods to determine the complete distributions of PWM scores and ratios between allele scores, facilitating assignment of statistical significance to rSNP effects. We tested our method on 41 experimentally verified rSNPs, correctly predicting the disrupted TF in 28 cases. We also analysed 146 disease associated SNPs from The Catalog of Published Genome-Wide Association Studies [2] with no known functional impact in an attempt to identify candidate rSNPs. We predicted 11 SNPs to be significantly disrupting the binding of a TF. Of these 11, 8 had previous evidence of the TF being associated with the disease in the literature. In addition, genes associated with each of the diseases were enriched for binding sites of the predicted disrupted TF in 5 cases. These results demonstrate that is-rSNP is suitable for high-throughput screening of SNPs for candidate

rSNPs. This is a useful and important tool in the interpretation of GWAS.

Author details

¹Department of Computer Science and Software Engineering, The University of Melbourne, Victoria, 3010, Australia. ²NICTA, Victoria Research Lab, The University of Melbourne, Victoria, 3010, Australia. ³Department of Biochemistry and Molecular Biology, The University of Melbourne, Victoria, 3010, Australia. ⁴Ian Potter Centre for Cancer Genomics and Predictive Medicine, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia. ⁵Bioinformatics and Systems Integration, The Blood and DNA Profiling Facility, Baker IDI heart and diabetes Institute, 75 Commercial Rd, Prahran, 3004, Victoria, Australia.

Published: 7 December 2010

References

1. Geoff Macintyre, James Bailey, Izhak Haviv, Adam Kowalczyk: **is-rSNP: a novel technique for in silico regulatory SNP detection.** *Bioinformatics* 2010, **26**(18):i524-i530.
2. Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: **A Catalog of Published Genome-Wide Association Studies.** 2010 [http://www.genome.gov/gwastudies].

doi:10.1186/1471-2105-11-S10-O7

Cite this article as: Macintyre et al.: is-rSNP: a novel technique for in silico regulatory SNP detection. *BMC Bioinformatics* 2010 **11**(Suppl 10):O7.

**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



* Correspondence: gmaci@csse.unimelb.edu.au

¹Department of Computer Science and Software Engineering, The University of Melbourne, Victoria, 3010, Australia

Full list of author information is available at the end of the article