

Lee-Ann Wood
330086
MSc (Med). Human Genetics
School of Pathology
Faculty of Science
University of the Witwatersrand

Supervisor: Professor Michèle Ramsay
Head of Molecular Genetics and Professor in the Division of Human genetics

Computational identification of synonymous SNPs in the human genome and their potential role in disease

ABSTRACT

The potential phenotypic effects of synonymous SNPs (sSNPs) have long been overlooked. Although several sSNPs are no longer thought to be silent, no one has identified which sSNPs may contribute to phenotypic variation on a genome-wide scale. sSNPs that cause a change in codon-usage frequency or mRNA secondary structures may alter translational and protein folding kinetics. In addition, sSNPs that alter splice-site consensus sequences may cause aberrant splicing, which could change the protein product. A sSNP that contributes to any of these molecular mechanisms may thus alter protein structure and function. To computationally identify sSNPs with a potential impact, SynSNP was created. SynSNP is a text-based tool written in Python. All sSNPs published within dbSNP are first identified. SynSNP uses established bioinformatics tools to determine which of the sSNPs may potentially result in a molecular effect. The potentially functional sSNPs are then assessed to determine whether any have previously been associated with a trait or disease in genome-wide association studies (GWAS) and/or occur within genes known to be associated with disease in OMIM (Online Mendelian Inheritance in Man). Of the 90,102 identified sSNPs, 21,086 (23.4%) were predicted to potentially have a functional impact, through one or more of the three molecular mechanisms investigated. Of the sSNPs predicted to potentially have a functional impact, 14 (0.07%) had previously been associated with a trait or disease in GWAS. A subset of 4,057 (19.2%) of the potentially functional sSNPs were within genes known to be associated with disease in OMIM. Only six (0.03%) of the potentially functional sSNPs had previously been associated with a trait or disease in GWAS and occurred within genes known to be associated with disease in OMIM. SynSNP could be developed further to aid the discovery of more sSNPs with a potential functional impact. A significant proportion of sSNPs may have a functional impact and their potential role in disease should therefore not be underestimated or neglected.