Methods

TAMAL is designed to be interactive, so that in addition to displaying suggested SNPs, the researcher can dynamically filter the results based on any of the application’s controls. On the left panel in Figure 1, the user inputs the standard gene name for a single gene or uploads a list of genes. The standard gene name is generally that approved by HUGO (http://www.gene.ucl.ac.uk/cgi-bin/nomenclature/searchgenes.pl), e.g. COMT for catechol-O-methyltransferase. All genomic locations are per the hg16 UCSC build.

The middle panel shows the result of querying the TAMAL database for the gene(s) input by the user. Optionally, the user can also limit the search to the most 5′- and 3′-extent of the gene or extend the search by a specified number of bases in either direction (20 000 bases by default). The SNP set is limited to those with evidence of variation in any of the major SNP databases (dbSNP, HapMap, Perlegen and Affymetrix).

The right panel lists sets of criteria that can be used to filter the set of SNPs according to flexible criteria. At the top, the user can select the Gabriel method (Gabriel et al., 2002) or TAGGER method of selecting haplotype tag SNPs from any or all of the four HapMap ancestry groups (The International HapMap Consortium, 2003) as determined by HaploView (Barrett et al., 2004). It is important to note that some genomic regions may not be amenable to this approach (Wall and Pritchard 2003a,b). At the middle of the right panel, the user can select SNPs that lead to non-synonymous or synonymous amino acid changes augmented with in silico prediction of functionality in silico (Karchin et al., 2005) or alter an intronic splice site. At the bottom, the user can select SNPs that occur in certain types of genomic features—SNPs that are in a predicted promoter (Wall and Pritchard 2003a,b) or predicted transfactor binding site (TRANSFAC v6.0, http://www.gene-regulation.com), along with SNPs that are in regions with conservation scores ≥99th percentile genome-wide for human–chimp–rat–mouse–chicken alignment via a hidden Markov model (Stormp and Haussler, 2003).

The user can inspect the choice of SNPs by clicking on the down arrow next to a gene in the middle panel. This opens the UCSC genome browser in the entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org
a separate window (inset in Fig. 1) so the user can inspect the SNP coverage and ensure that the SNPs selected are a reasonable subset of all those potentially available. Finally, at the lower edge of the middle panel users can download the results into an EXCEL file (commonly used by researchers) or in XML format (for exchange with other applications).

TAMAL is provided as a good faith effort to assist the human genetics community. No such tool should be considered as a foolproof ‘black box’. There are some genes that will be difficult to study with typical SNP methods, and there are additional databases for some genes that should be consulted (e.g. for genotyping members of the large CYP gene family). Nonetheless, provided that users remain cognizant of its limitations, TAMAL can greatly assist with rational SNP selection.

We will endeavor to update TAMAL on a quarterly basis to incorporate updates to the primary databases as well as new features.

ACKNOWLEDGEMENTS

We thank the Carolina Center for Exploratory Genetic Analysis for computational support (P20RR20751), and the Informatics and Visualization Laboratory (http://www.ils.unc.edu/bmh/vlab) at the School of Information and Library Science for hosting this service. Funding to pay the Open Access publication charges was provided by the University of North Carolina at Chapel Hill’s Open Access Publishing Fund.

Conflict of Interest: none declared.

REFERENCES