

Identifying Candidate Coding Region Single Nucleotide Polymorphisms (cSNPs) and Alternative Splice Variants in Human Genome Using Assembled Expressed Sequence Tags (ESTs)

Kavita Garg, Deborah Nickerson and Phil Green
kavitag@u.washington.edu

Dept of Molecular Biotechnology
University of Washington
Seattle

Databases of human ESTs derived from different individuals, tissues and developmental stages are rich resources for finding SNPs and alternative splice forms. Using assembled ESTs (C. Wilson and P. Green, unpub) from 50 different cDNA libraries, we have identified contigs that represent the complete coding sequences of 850 known human genes, and have scanned them for high quality substitutions (Garg et al. 1999). We report the analysis and characteristics of candidate cSNPs found in coding regions of 165 of these genes. Our analysis has identified 201 candidate cSNPs, of which 87 are predicted to lead to amino acid changes.

We are currently analyzing assembled ESTs (Green and Ewing 2000) to identify and classify the forms of alternative splicing in human genes. Preliminary results from these analyses will also be presented.

References

Garg K., P. Green, and D.A. Nickerson. 1999. Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. *Genome Research*. 9:1087-1092.

Ewing B, P. Green. 2000. Analysis of expressed sequence tags indicates 35,000 human genes. *Nature Genetics* 25: 232-234.