Title:
Ensembl Variant Effect Predictor - flexible and consistent molecular consequence prediction

Authors:
Irina M Armean, Laurent Gil, Diana Lemos, Andrew Parton,Helen Schuilenburg, Anja Thormann, Sarah E Hunt, Paul Flicek, Fiona Cunningham

European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom

Molecular consequence prediction is the cornerstone of variant interpretation and consistent reporting is essential in both basic and clinical research settings. The Ensembl Variant Effect Predictor (VEP) predicts the molecular consequence of variants on transcripts, and regulatory features, alongside reporting scores from a range of variant assessment tools and information from reference datasets e.g. allele frequencies, phenotype associations and citations. We have recently made enhancements to support more standardised variant consequence reporting.

Ensembl VEP can annotate variants using both the Ensembl/GENCODE and RefSeq transcript sets, which differ slightly as they are created using slightly different annotation strategies. EMBL-EBI and NCBI are now collaborating to create the MANE (Matched Annotation from NCBI and EMBL-EBI) Select transcript set, which will include a representative transcript for each protein coding gene. This transcript set, which will support consistent variant reporting across different tools and transcript sets, is now available in Ensembl VEP. As we will demonstrate here, it remains important to consider the effect of a variant on all relevant transcripts.

The sharing and comparison of variant annotations may also be confounded by insertions and deletions within repeated sequence, which can be represented as multiple equivalent genomic changes. We have updated Ensembl VEP to optionally normalise variants to their most 3’ representation before analysis. This normalises variant consequence annotations in a biologically appropriate manner and is in accord with Human Genome Variation Society (HGVS) nomenclature.