# Large scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing

Picardi E[1], Horner DS[2], Chiara M[2], Schiavon R[3], Valle G[3], Pesole G[1,4]

## Motivation

RNA editing is a widespread post-transcriptional molecular phenomenon that can increase proteomic diversity, by modifying the sequence of completely or partially non-functional primary transcripts, through a variety of mechanistically and evolutionarily unrelated pathways [1]. Editing by base substitution has been investigated in both animals and plants and recently identified in viruses [2, 3]. Post-transcriptional alterations due to editing are currently detected by the classical Sanger sequencing and comparing the genomic locus with the corresponding cDNA. Such conventional strategy, however, is time-consuming and effectively prevents a genome wide identification of RNA editing. The tissue-specific assessment is also precluded. In contrast, the high-throughput RNA-Seq approach allows the generation of a comprehensive landscape of RNA editing at the genome level. Millions of short reads from Solexa/Illumina GA and ABI SOLiD platforms have been used to de novo investigate the editing pattern in mitochondria of Vitis vinifera [4].

## Methods

More than 205 million reads from four grapevine (cultivar PN40024) tissues (root, leaf, callus and stem) were sequenced with the Solexa/Illumina GA technology and subsequently mapped onto the corresponding mitochondrial (mt) genome using the program PASS. In addition, 328 millions of ABI SOLiD reads from leaf and root tissues were also mapped onto the color space version of the mt genome using the same software. Individual alignments, however, were parsed by custom scripts in order to distinguish color space changes due to sequencing errors from genuine polymorphisms. For each mt genome position falling in a known annotation we generated the distribution of mapping nucleotides, filtering out bases according to a fixed quality threshold. Then, statistically significant C-to-U conversions were detected by applying the Fisher's exact test by comparing the observed and expected C and U occurrences in the aligned reads. Expected C and U occurrences were estimated using sequencing error rates.

[1] Dipartimento di Biochimica e Biologia Molecolare "E. Quagliariello", Università degli Studi di Bari, Bari [2] Dipartimento di Scienze Biomolecolari e Biotecnologie, Università degli Studi di Milano, Milano [3] CRIBI, Università degli Studi di Padova, Padova [4] Istituto Tecnologie Biomediche del Consiglio Nazionale delle Ricerche, Bari

**Results**

Grapevine RNA-Seq data from Solexa/Illumina and SOLiD platforms provide a significant support (at 0.05 confidence level) for 401 C-to-U modifications in grapevine coding regions. Additional 44 editing alterations, instead, involve non-coding RNAs (tRNAs and group II introns). Interestingly, over 90% of significantly detected edited sites are conserved at the same positions in other plant species. Moreover, of these editing events Solexa/Illumina supports the 24.6% whereas SOLiD data supports the 75.4%. Indeed, Solexa/Illumina and SOLiD platforms show different characteristics with respect to the specific issue of large-scale editing analysis even though a combined approach may reduce false positive detections. Finally, 76% of all C-to-U conversions in coding genes represent partial RNA editing events and 28% of them were shown to be significantly tissue specific. Analyzing all 401 detected sites, we were not able to identify a strict consensus motif for sequencing surrounding RNA editing changes. However, it is evident a bias towards pyrimidines at positions -2 and -1 and a bias towards purines at position +1, as demonstrated also for other plants mt genomes.

**Contact e-mail**

e.picardi@biologia.uniba.it

**Supplementary information**

References

[1] Brennicke, A., Marchfelder, A. and Binder, S. (1999) FEMS microbiology reviews, 23, 297-316 [2] Gott, J.M. and Emeson, R.B. (2000) Annual review of genetics, 34, 499-531. [3] Casey, J.L. (2006) Current topics in microbiology and immunology, 307, 67-89 [4] Picardi, E. et al. (2010) Nucleic Acids Research, in press