



## **ABSTRACT**

### *“Selection on noncoding DNA in the Brassicaceae”*

**Dr Annabelle Haudry**

**Laboratoire de Biometrie et Biologie Evolutive**

**Universite Lyon 1 - UMR CNRS 5558**

**Villeurbanne cedex**

**FRANCE**

Despite the central role of noncoding DNA in gene regulation and evolution, our understanding of the genomic extent and nature of selection on plant noncoding regions remains limited. Taking advantage of newly sequenced genomes in the Brassicaceae, we combined comparative genomics and population genomics approaches to analyze patterns of molecular evolution and the strength of selection of noncoding DNA. We conducted whole genome alignments of nine Brassicaceae species, including two new genome assemblies from the DOE Joint Genome Institute (*Capsella rubella* and *Eutrema halophila*) and de novo Illumina assemblies of three additional species by the VEGI project (*Aethionema arabicum*, *Sisymbrium irio*, and *Leavenworthia alabamica*). Using a systematic computational approach we identified over 90,000 conserved noncoding regions (CNSs) that show a degree of conservation across species similar to average coding exons and no evidence of expression or coding potential. Gene ontology analysis revealed a significant enrichment of CNSs near transcription factors and CNSs were found to be enriched in known transcription factor binding sites. These results support the potential regulatory function of CNSs. Population genomic analyses in *A. thaliana* and *C. grandiflora* showed reduced levels of diversity and excess of rare variants for CNSs compared to neutral expectation, confirming that most CNSs evolve under purifying selection. Our analysis indicates the strength of selection acting on plant noncoding regions, highlighting the likely importance of these regions for functional variation and adaptive evolution.