

SPP 2349 (GEvol): Genomic basis of evolutionary innovations.

John Parsch, Munich
Erich Bornberg-Bauer, Münster

Emergence, loss and regulation of *de novo* genes

Abstract

In recent years, a major shift has occurred in understanding how new proteins arise and gain function. Earlier research assumed that small, adaptive changes in duplicates of established genes led to the evolution of new proteins with new functions. However, recent research showed that new protein-encoding genes can arise "de novo" from previously non-coding DNA. Although some mechanisms underlying de novo gene emergence are now understood, no coherent picture of how and how often novel protein coding genes arise, gain function and spread through a population is available. The aim of this project is to characterise the short and long-term evolutionary dynamics of de novo gene emergence, loss and regulation. This will be done by sequencing genomes and transcriptomes of multiple *Drosophila* species spanning 3–40 million years of evolutionary divergence, including multiple individual genomes per species. This will allow us to precisely pin down the most recent events, such as the emergence of regulatory sequences that may have triggered the expression of an intronic or intergenic region. Including longer time scales will allow us to understand how often protein coding genes are created de novo and how often they are lost – most likely because they had not assumed a sufficiently important function to be retained by selection. Bioinformatic, population genetic and experimental functional genetic analyses will be performed to expand our understanding of how genomes change over time to give rise to new molecular and phenotypic traits.