The "evol" is in the details: a rummage-region model for the origins of lineage-specific elements via gene duplication, relocation, and regional rearrangement in *Neurospora crassa*

Zheng Wang[1], Yen-Wen Wang[1], Takao Kasuga[2], Hayley Hassler[1], Francesc Lopez-Giraldez[3], Caihong Dong[4], Oded Yarden[5], and Jeffrey Townsend[6]

[1]Yale University School of Public Health
[2]University of California Davis
[3]Yale School of Medicine
[4]Institute of Microbiology Chinese Academy of Sciences
[5]Hebrew University of Jerusalem Robert H Smith Faculty of Agriculture Food and Environment
[6]Yale University

April 27, 2023

Abstract

The origin of new genes has long been a central interest of evolutionary biologists. However, novelty evades reconstruction by the classical tools of evolutionary modeling. This evasion of insight from deep ancestral investigation necessitates intensive study of model species within well-sampled, recently diversified clades. The model *Neurospora* species—which lack recent gene duplications yet harbor clusters of lineage-specific genes (LSGs) adjacent to the telomeres—constitute comprehensively characterized organisms apt for studying the evolution of LSGs. Using gene syntenies, we documented that 78% of Neurospora LSGs clusters accompany large non-coding regions, frequent gene duplications and relocation, or regional rearrangements. Ancestral status of the LSG *mas-1* and its neighbors was investigated in detail, and we identified sequence conservation among syntenic non-coding regions that suggests that it arose from an ancient copy of a lysophospholipase precursor that is ubiquitous in lineages of the Sordariomycetes. High resistance to polyoxin D of the *mas-1* mutant demonstrates that the gene exhibits a role in cell-wall integrity and cellular sensitivity to antifungal toxins. To perform a broader investigation of the function of LSGs, we assembled transcriptomics data from 68 experimental data points and identified co-regulatory modules using Weighted Gene Correlation Network Analysis. This analysis revealed no essential roles for LSGs in known regulatory machinery. Our discoveries illuminate a "rummage region" in the *N. crassa* genome that enables some novel elements and new functions to arise via gene duplication and relocation or invasion of genetic materials, followed by fast mutation and recombination facilitated by tandem repeats and unconstrained non-coding sequences.

Hosted file

`New version of orphan_gene_Mol_Ecol.docx` available at https://authorea.com/users/611982/articles/640175-the-evol-is-in-the-details-a-rummage-region-model-for-the-origins-of-lineage-specific-elements-via-gene-duplication-relocation-and-regional-rearrangement-in-neurospora-crassa