A New Gene Becomes Old Again

July 15, 2013

GMR Deyter

The conservation of proteins that fulfill critical cellular processes is well appreciated. As an example, microtubules are involved in a variety of processes in different cell types, ranging from vesicle trafficking in neurons to cell movement in cilia-containing cells. The tubulin proteins that form microtubules are among the most conserved proteins throughout evolution, highlighting their critical roles in multiple processes. Tubulin is considered an 'old' gene, encoded in the DNA of early eukaryotic organisms and essential for organismal survival. However, it has become apparent that 'new' genes that exist in only one or a few species can quickly become essential for survival. The steps by which 'new' genes acquire their essential function through rapid evolution, however, remain enigmatic. In a recent manuscript published by the Malik Lab at FHCRC, lead author Benjamin Ross and his co-authors (including scientists from the University of Connecticut and Ludwig Maximilians University of Munich, have explored how a new and essential gene named *Umbrea* obtained centromere localization and function in the fruit fly *Drosophila*.

A major mechanism of new gene formation is gene duplication, and *Umbrea* arose from the duplication of the*HP1B* (*Heterochromatin Protein 1B*) gene. To compare the functions of the two proteins, the authors tagged HP1B and Umbrea with green fluorescent protein (GFP) to visualize the proteins in cells. The researches found that GFP-HP1B localizes to heterochromatin in *Drosophila melanogaster* (*D. melanogaster*), a pattern distinct from the centromere localization of GFP-Umbrea. In addition, only Umbrea, but not HP1B, is an essential gene in *D. melanogaster*, indicating critical changes in Umbrea protein structure and function occurring since its birth in the genome. The authors observed chromosome segregation defects in cells depleted of Umbrea suggesting that Umbrea's centromere localization corresponds to an essential function of Umbrea during mitosis. But when and how did Umbrea obtain these burgeoning features?

Ross *et al.* sequenced the *Umbrea* locus from 32 *Drosophila*species and found *Umbrea* in only 20 species. This phylogenetic analysis dated its origin to 12 to 15 million years ago. The sequencing data also revealed interesting changes that occurred to the original *Umbrea* locus during evolution. *HP1B* and the most ancestral *Umbrea* encode proteins that have both chromodomain (CD) and chromoshadow domains (CSD). CDs bind to H3K9 trimethylated histones, consistent with

a role of HP1B in binding this histone modification that is enriched in heterochromatin. Although most existing *Umbrea* genes have lost their CDs, this genetic alteration was not enough to target Umbrea to centromeres given the authors' data showing that a HP1B mutant lacking the CD lost heterochromatin localization but did not target to centromeres. Furthermore, fusing the HP1B CD to Umbrea perturbed Umbrea's centromere localization. These results led Ross and colleagues to speculate that additional genetic alteration had to occur to *Umbrea* for it to gain an essential function at centromeres.

The researchers found that the CSD of Umbrea from *D. melanogaster* localized to centromeres, indicating a critical role for the CSD in centromere targeting. Since CSDs are involved in protein-protein interactions, the authors tested whether changes in the CSD of Umbrea altered its protein-binding network. An amino acid alignment of the CSD of HP1B and Umbrea from a variety of *Drosophila* species revealed the presence of residue changes that were anticipated to alter the binding partners of Umbrea compared to HP1B. This expectation was confirmed when novel Umbrea binding partners were identified by mass spectrometrical analysis of Umbrea purified from insect cells (including bona fide centromere proteins).

Additionally, Ross *et al.* reveal that short tail sequences that flank the CSD of Umbrea participate in centromere localization of Umbrea. The authors suggest that these amino acids might interact with DNA, and the intriguing possibility that these rapidly evolving tails have co-evolved with the divergence of centromeric DNA sequence between species to give Umbrea species-specific centromere localization. Evolutionary genetics, similar to Umbrea itself, is a burgeoning field that is proving to be essential in discovering the genetic innovations that have shaped nearly every aspect of biology.

Ross BD, Rosin L, Thomae AW, Hiatt MA, Vermaak D, de la Cruz AF, Imhof A, Mellone BG, Malik HS. 2013. Stepwise evolution of essential centromere function in a Drosophila neogene. *Science* 340(6137):1211-4.

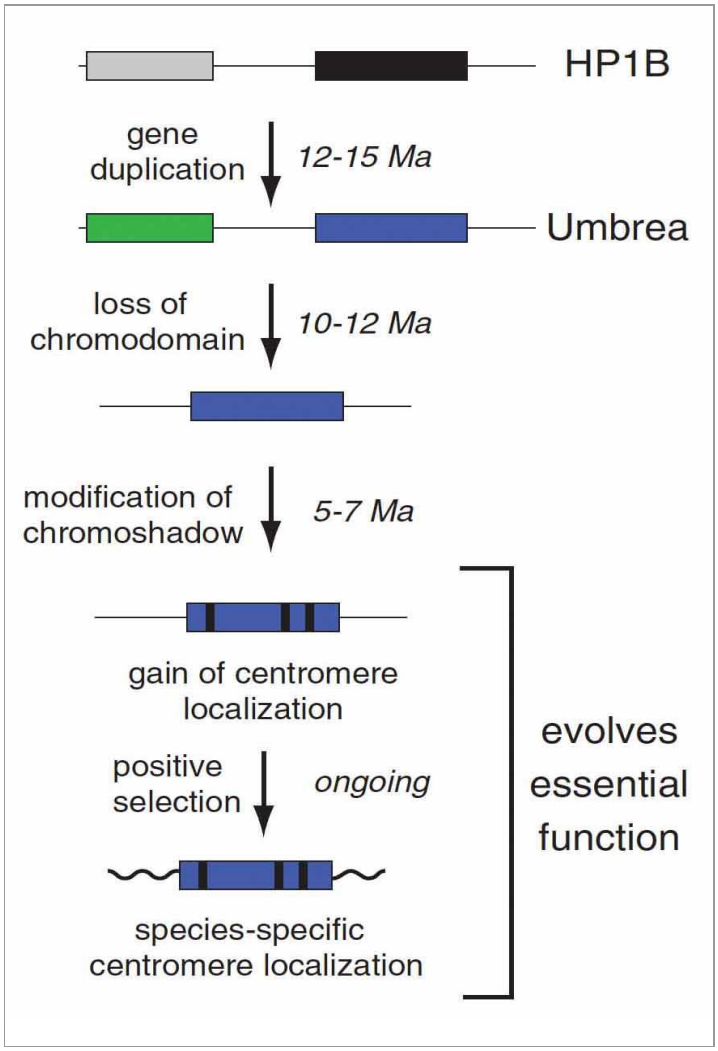

Image adapted from the manuscript

The genetic steps by which Umbrea acquired essential centromere function through time. Umbrea arose from a HP1B gene duplication occurring around the indicated time (Ma=million years ago). Modification of the chromoshadow domain subsequent to loss of the chromodomain gave Umbrea centromere localization.