Heredity (2004) 93, 122–123

www.nature.com/hdy

NEWS AND COMMENTARY

Genome complexity

Adaptive evolution or genetic drift? Does genome complexity produce organismal complexity?

RB Phillips

Heredity (2004) **93,** 122–123, advance online publication, 12 May 2004; doi:10.1038/sj.hdy.6800494

In their recent *Science* paper, Lynch and Conerly, 2003 argue that in the transition from prokaryotes to multicellular eukaryotes, population sizes decreased dramatically as the size of organisms increased. This magnified the power of random genetic drift and allowed the proliferation of genome features that would have been eliminated by purifying selection in larger populations. Once these features were in place, they provided the raw material for evolution of phenotypic complexity by natural selection.

In support of the first part of their hypothesis, the authors present convincing data showing that prokaryotes have larger effective population sizes than do most eukaryotes. In fact, one of the few supposed exceptions to this trend is probably not an exception after all. The ciliate *Tetrahymena thermophilia* is listed as having a very large population size, more in keeping with prokaryotes, but the reference cited ignores the fact that this 'species' is in reality a 'species complex' composed of reproductively isolated syngens.

The new genomic complexity found in higher organisms includes introns, mobile genetic elements, and an increase in duplicate genes. Mobile elements are mainly found in genome sizes above 100 MB, and larger introns are found in areas of the genome with low recombination rates (Carvalho and Clark, 1999). These observations support their idea that these features might be retained passively in larger genomes in response to a reduction in purifying selection. The suggestion that the half-life of duplicate genes might increase with genome size in response to decreasing effective population size is also plausible. If so, genetic drift has had a much greater role in adaptive evolution in complex genomes than has been envisioned by advocates of neutral theory.

The underlying assumption in Lynch and Conery's report is that this genomic complexity provides the raw material for organismic complexity, although the authors acknowledge that the two are often not very well correlated. For example, the fact that subfunctionalization of genes is much more likely than neofunctionalization means that possession of extra copies of genes does not usually lead to evolutionary innovation.

Although the genomes of eukaryotes are generally larger than those of prokaryotes, genome size is not correlated with organismal complexity. Unicellular eukaryotes have genome sizes that vary over 200 000-fold, with the genome of the *Amoeba* being about 200 times greater than that of humans (Gregory, 2001). The number of genes in the sequenced organisms, in general, shows a gradual increase in organismal complexity, from *Escherichia coli* with about 4300 genes, and yeast with 6000, to *Drosophila* with 15 000 and humans with 30 000. However, *Caenorhabditis elegans* has 21 000 genes and is morphologically less complex than *Drosophila*. The estimated number of genes for the ciliate *Paramecium tetraurelia* is similar to that for humans (McGrath and Katz, 2004). There are many examples in higher plants of polyploids with large num-bers of duplicate genes, but no more organismic complexity than in related diploids.

Genomic turnover is not related to organismic change either. Sequence divergence is substantial between some of the *Tetrahymena* 'syngens' that appear identical. Sequencing of a second worm has revealed that there is a three-fold larger sequence divergence between the two worms (*C. brigissae* and *C. elegans*) than between humans and mice (Blaxter, 2003).

So, what does determine organismic complexity? Levine and Tjian, 2003 argue that organismic complexity correlates with an increase in the ratio and number of transcription factors per gene. The yeast genome has 300 transcription factors, but there are 1000 in *Drosophila* and possibly 3000 in humans. The promoter regions in higher organisms are much larger, and there appears to be a much greater variety of protein complexes that interact with these regulatory DNAs, which help to provide the tissue specificity of gene expression found in multicellular organisms.

These differences in the number of regulatory protein complexes between yeast and higher organisms may explain the findings in another recent paper by Yang and Li (2003). The authors reason that protein complexity, defined as the number of subunits in a protein (n), might explain whether duplicate genes are retained, since duplication of one subunit might cause a dosage imbalance among the subunits of a protein. They found that the proportion (P) of unduplicated genes increased with the number of subunits in a protein. However, P was higher for both monomers and multimers in yeast, but low in humans, and the family size of genes was also significantly higher in humans compared with yeast. These results suggest that organismal complexity is a stronger determinant of gene duplicability than is protein complexity, and are consistent with the analysis of Lynch and Conerly (2003) as well as that of Levine and Tjian (2003). The duplicate genes would have a higher probability of being retained in higher organisms, and this could lead to duplicate sets of regulatory multimers that acquire tissue-specific functions.

The idea that evolution of regulatory genes might explain the disconnection between genome change and organismic change was suggested about 30 years ago by Allan Wilson and collaborators (Cherty *et al*, 1978). This was prompted by the discovery of the large genome differences among morphologically similar frog species and the small nucleotide difference between chimpanzees and humans. Thus, there are widespread examples of structural stasis in the face of substantial genomic change from prokaryotes to unicellular eukaryotes to higher organisms including vertebrates.

Multicellular organisms have evolved multiple times in prokaryotes and eukaryotes (Kaiser, 2001) and in some cases related unicellular and multicellular relatives live in the same environment. It has been hypothesized that advantages in feeding and dispersion may have driven the evolution of these new forms. The genomes of several of these species pairs are currently being sequenced, and some of these species can be maintained in the laboratory, so they will be excellent models for in-

vestigation of the forces involved in evolution of organismic complexity.

In response to environmental change on our planet, most species have become extinct, some have retained their structural integrity and a few have evolved greater organismic complexity. Understanding how multicellular organisms evolved will in turn require a deeper understanding of the organization of this complexity.

RB Philips is at the Department of Biological Sciences, Washington State University, Vancouver, WA 98686-9600, USA.

e-mail: philipsr@vancover.wsu.edu

Lynch M, Conerly JS (2003). *Science* **302**: 1401–1404.
Carvalho AB, Clark AG (1999). *Nature* **401**: 344.
Gregory TR (2001). *Biol Rev* **76**: 65–101.
McGrath CL, Katz LA. (2004). *Trends Ecol Evol* **19**: 32–38.
Blaxter M (2003). *Nature* **426**: 395–396.
Levine M, Tjian R (2003). *Nature* **424**: 147–151.
Yang, Li (2003). *Proc Natl Acad Sci USA* **100**: 15661–15665.
Cherty LM, Case SM, Wilson AC (1978). *Science* **200**: 209–211.
Kaiser D (2001). *Ann Rev Genet* **35**: 103–123.