Evolution of the yeast species concept in the age of sequencing

Marc-André Lachance
UWO, University of Western Ontario, London, Ontario, Canada N6A5B

Although the strain is the unit of interest to microbial culture collections, the species is the fundamental unit of biodiversity. The yeast species concept has changed considerably in past decades, based successively on morphology, physiology, whole genome similarity, and most recently sequence divergence. As an objective concept founded on membership in a reproductive community is seldom attainable, proxy concepts are necessary. After reviewing some of the principal metaphysical elements of the yeast species concept, I shall focus on the criterion that is currently applied by most researchers involved in the description of new yeast species, namely the extent of divergence in sequences such as the D1/D2 variable domains of the LSU rRNA gene. Kurtzman and Robnett (1998) observed that members of well-defined ascomycetous yeast species rarely differ by more than three substitutions in D1/D2 sequence and that representatives of different species tend to show at least 1% divergence in that sequence. This empirical observation has not been discussed in a theoretical framework. In particular, the impact of sampling effort on the amount of sequence diversity allowed within a species has been largely overlooked. Many yeast systematists might agree that a nomenclatural yeast species should ideally reflect the biological species concept, as it corresponds to a real evolutionary unit. Darwin, in 1859, stated that the task of systematists is to recognize species from mere varieties. Hennig (1962) provided a solution to this challenge by distinguishing between two forms of variation, namely phylogenetic variation, which accompanies speciation, and tokogenetic variation, known in evolutionary genetics as polymorphism. Templeton et al. (1992) have provided a statistical test that allows the erection of meaningful boundaries between species in the absence of more objective criteria. I shall discuss the use of haplotype parsimony networks in interpreting sequence variation to delineate species. My examples will come from analyses of ITS and D1/D2 sequences of well-sampled but polymorphic species such as *Candida azyma*, *Candida apicola*, *Metschnikowia agaves*, and *Starmerella bombicola*. I shall show that poor sampling can lead to the description of multiple species when in fact the strains in question should be assigned to a single species.

Key words: Species, Yeast, Polymorphism, DNA sequencing, Parsimony networks