(Deutsch unten)

Microsporidia are a group of parasites that infect a wide range of species, many of which play important roles in agriculture and human disease. Approximately 1,400 species of microsporidia have been described. In which, At least 14 microsporidian species have been confirmed to cause dangerous infectious diseases in both immunocompromised and immunocompetent humans.

Originally microsporidia were classified as fungi by Naegeli (1857). Later, they were placed then within the Archezoa kingdom with some strong evidences based on some detailed morphological researches. The placement of microsporidia as an early branching eukaryote however remained a topic for debates. Recently, many molecular phylogeny studies revived the original association of microsporidia with fungi. The exact microsporidia-fungi relationship is unresolved yet.

The "Microsporidia-early" hypothesis can be explained by their lack of several cellular components as well as by their reduced genomes and metabolism. Being obligate intracellular parasites, microsporidia successfully reduced their genome sizes, down to the range of bacteria. They are thought to retain only genes essential for their survival and development. Moreover, microsporidia are obligatorily dependent on many primary metabolites from the hosts due to the absence of several key pathways, such as the citric acid cycle, oxidative phosphorylation, or the *de novo* biosynthesis of nucleotides. The ancestral genome of microsporidia, however, was supposed to be more complex. Consequently, the reduction in microsporidian genomes and metabolism would be the result of secondary processes.

Microsporidia are becoming model organisms for studying minimal eukaryotic genomes and obligate endoparasitic lifestyles in eukaryote. There is little knowledge regarding the evolution of the microsporidia and their metabolisms as different factors make their experimental study challenging. Comprehensive comparative analyses of the microsporidia are thus needed.

We investigated the evolutionary history of the contemporary microsporidia though the reconstruction and investigation of their last common ancestor (LCA). To enable the intuitive data analysis of the microsporidian phylogenetic profiles, we have developed PhyloProfile, a program that facilitates a dynamic visualization and exploration of the presence-absence patterns of microsporidian genes integrated with other additional information layers. Additionally, we present HamFAS as a new approach for the reliable functional annotation of proteins. HamFAS integrates a targeted ortholog search based on the HaMStR algorithm with a weighted assessment of feature architecture similarities (FAS) between orthologs. Comparing with two state-of-the-art annotation tools KAAS and BlastKOALA using a manually curated yeast protein set, HamFAS yielded the best precision (98.5%) and a higher sensitivity with almost two times more proteins annotated with HamFAS than with the other tools.

We furthermore reconstruct a robust phylogeny of microsporidia to identify the placement of the microsporidia in the eukaryotic tree of life using the data available from eleven microsporidian genome sequencing projects. We successfully identified a microsporidian core set comprising of 80 one-to-one orthologous groups, which were used as a basic for the phylogeny reconstruction. Our reconstructed phylogenetic tree from a diverse set of taxa, including 48 fungi representing the full fungal diversity as well as other representative species for both unikonts and bikonts, strongly supported the hypothesis that microsporidia form the early branch of fungal clade. The analyzed data explained this microsporidia-fungi relationship significantly better than any other hypotheses.

Our phylogenetic profile study clarified the dynamic evolutionary history of microsporidian genomes. Between 2% of the proteins in the compact microsporidia *Encephalitozoon intestinalis* and up to 49% of the proteins of *Edhazardia aedis* are private for individual microsporidian species. Those private proteins are useful to investigate adaptations of each microsporidium to their host. On the other hand, proteins that are found in all microsporidia are expected to be essential and evolutionary old. We found that 94% of the microsporidian LCA proteins could be tracked back to the last eukaryotic common ancestor and further 3% of the LCA proteins were shared in the common ancestor with fungi. Only 3% of the LCA proteins are microsporidia specific, which are potentially important for the obligate parasitic lifestyle of the microsporidia.

The functional annotation and metabolic pathway analysis of the microsporidian LCA protein set gave us more insight into the adaptation of the microsporidia to their parasitic lifestyle and the origin of the microsporidian genome reduction. The presence of E1 and E3 components of the pyruvate dehydrogenase complex and the mitochondrial hsp70 protein support an ancestral presence of mitochondria in the ancestral microsporidia. In addition, several ancient proteins that complement gapped metabolic pathways were found in the microsporidian LCA. They suggested a more complex genome and metabolism in the LCA. However, the microsporidian LCA still lacked many main pathways such as the TCA cycle for effective energy production, and key enzymes that are required for *in vivo* synthesis of critical metabolites like purines and pyrimidines. In summary, we find that the parasitic lifestyle and the genome reduction already occurred in the microsporidian LCA. This ancestral state was followed by further losses and gains during the evolution of each individual microsporidian lineage.

Mikrosporidien sind sporenbildende Parasiten, die verschiedene Organismen infizieren. Die Mikrosporidiose beeinträchtigt nicht nur die Agrarökonomie, sondern führt auch zu humanmedizinischen Krankheitsbildern. Etwa 1.400 Mikrosporidien wurden bislang beschrieben. Indem sind mindestens 14 Mikrosporidien bekannt, welche gefährliche infektiöse Krankheiten sowohl bei immunkompromittierten als auch bei immunkompetenten Menschen verursachen.

Mikrosporidien wurden erstmals 1857 von Naegeli als Pilz klassifiziert. Einige detaillierte morphologische Untersuchungen gaben später starken Evidenzen für die Platzierung der Mikrosporidien innerhalb des Archezoa-Reiches. Trotz dieser Evidenzen wurde die "Microsporidia-early" Hypothese in Frage gestellt. Vor Kurzem wurde die Verwandtschaft zwischen Pilzen und Mikrosporidien erstmals durch molekulare phylogenetische Studien unterstützt. Die exakte Position der Mikrosporidien bleibt aber im Speziesbaum der Pilze immer noch ungelöst.

Die "Microsporidia-early" Hypothese kann/konnte durch das Mangel vieler typischen eukaryotischen Komponenten als auch den reduzierten Genome der Mikrosporidien unterstützt. Als obligate intrazelluläre Parasiten verminderten die Mikrosporidien ihre Genome soweit das deren Größe im Bereich bakterieller Genome liegt. Die geringe Anzahl von protein-kodierenden Genen deutet darauf hin, dass die Genome der Mikrosporidien nur Gene enthalten die für ihr Überleben und ihre Entwicklung essentiell sind. Darüber hinaus fehlen den Mikrosporidien mehrere Stoffwechselwege, wie der Zitronensäurezyklus, die oxidative Phosphorylierung oder die *de novo* Biosynthese von Nukleotiden. Dies führt zu einer obligaten Abhängigkeit vom Wirt für viele primäre Metabolite. Das uncharakterisierte hsp70-Protein setzt jedoch ein komplexeres Genom des mikrosporidischen Vorfahren voraus. Folglich ist die Reduktion der mikrosporidischen Genome und Stoffwechselwege auf Sekundärprozesse zurückzuführen.

Mikrosporidien können als Modellorganismen für die Untersuchung eines minimalen eukaryotischen Genoms und obligater endoparasitischen Lebensweisen in Eukaryoten dienen. Dennoch gibt es bislang nur wenig Studien über die ihre Evolution und Metabolismus. Deshalb bedarf es einer umfassenden vergleichenden Analyse der Mikrosporidien.

Wir untersuchten daher die Entwicklungsgeschichte der heutigen Mikrosporidien durch die Rekonstruktion und Untersuchung ihres letzten gemeinsamen Vorfahrens (LCA). Für eine intuitive Datenanalyse der phylogenetischen Profile der Mikrosporidien entwickelten wir PhyloProfile, ein Programm für dynamische Visualisierung und Untersuchung von phylogenetischen Profilen. Außerdem präsentieren wir HamFAS, einen neuen Ansatz zur zuverlässigen funktionellen Annotation von Proteinen. HamFAS integriert eine gezielte Orthologen-Suche basierend auf dem HaMStR-Algorithmus mit einer gewichteten Bewertung von Feature Architektur Ähnlichkeiten (FAS) zwischen Orthologen. Wir benchmarkten die Performance von HamFAS mit einem manuell kuratierten, KO-annotierten, Hefeprotein-Set. Im Vergleich zu zwei State-of-the-Art Annotationstools KAAS und BlastKOALA, HamFAS erzielte die beste Genauigkeit (98,5%) und zeigte eine höhere Sensitivität, indem HamFAS annotierte im Durchschnitt 47,5% mehr Proteine als die anderen Tools.

Um die Mikrosporidien besser im eukaryotischen Stammbaum zu platzieren erstellten wir eine robuste Phylogenie der Mikrosporidien. Um dies zu erreichen, verwendeten wir die Daten von elf Mikrosporidien-Genom-Sequenzierungsprojekten. Wir identifizierten einen mikrosporidisches Geneset, welches aus 80 eins-zu-eins-Orthologen Gruppen besteht. Diese Gruppen dienten als Grundlage für die Phylogenie-Rekonstruktion. Unser rekonstruierter phylogenetischer Baum beinhaltet 48 Pilze, die die gesamte Pilzdiversität abbilden, sowie andere repräsentative Arten der Unikonten und Bikonten. Der resultierende Baum unterstützte die Hypothese, dass Mikrosporidien den frühesten Zweig der Pilzevolution bilden. Die analysierten Daten erklärten diese Mikrosporidien-Pilz-Verwandtschaft signifikant besser als alle anderen Hypothesen.

Unsere Studie der phylogenetischen Profile zeigte die dynamische Evolutionsgeschichte der mikrosporidischen Genomen. Je nach Mikrosporidium finden wir das zwischen 2% der Proteine bis hin zu 49% der Proteine nur in einer Art gefunden werden. Diese privaten Proteine deuten darauf hin, dass sie für die Wirtsanpassung benötigt werden. Im Gegensatz dazu stehen jene Proteine die in anderen Mikrosporidien gefunden werden und als essentiell und evolutionär alt angesehen werden. Wir finden, dass 94% der Proteinen des mikrosporidischen LCAs auf den letzten gemeinsamen Vorfahren aller Eukaryoten zurückverfolgt werden können und dass 3% der Proteine, mit dem gemeinsamen Vorfahren der Pilze geteilt werden. Nur 3% LCA-Proteine sind Mikrosporidien-spezifisch. Diese sind potentiell für die obligatorische parasitäre Lebensweise von Mikrosporidien zentral.

Die funktionelle Annotation und die Analyse der Stoffwechselwege des mikrosporidischen LCAs ermöglichte ein besseres Verständnis der Anpassung von Mikrosporidien an ihre parasitäre Lebensweise und den Ursprung ihrer Genom-Reduktion. Die Anwesenheit von E1, E3-Komponenten des Pyruvat-Dehydrogenase-Komplexes und des mitochondrialen hsp70-Proteins deuten darauf hin, dass die anzestralen Mikrosporidien Mitochondrien besaßen. Zusätzlich wurden mehrere alte Proteine im mikrosporidischen LCA gefunden, die einige Lücken Stoffwechselwege schließen können. Dies deutet auf ein komplexeres Genom und einen aufwändigeren Metabolismus im LCA hin als bislang vermutet. Dem mikrosporidischen LCA fehlen jedoch weiterhin viele primäre Stoffwechselwege oder Schlüsselenzyme, nehmen wir an, dass die parasitische Lebensweise bereits in der mikrosporidischen LCA vorkam. Die reduzierten Genome sind damit der anzestrale Zustand für die Mikrosporidien, welchem weitere Genverluste und Genzuwächse auf einzelnen mikrosporidischen Linien folgte.

Naegeli, K. 1857. "Über die neue Krankheit der Seidenraupe und verwandte Organismen." *Botanische Zeitung*, 1857, 760-761. Accessed 2018-03-25 20:33:39.