Ascomycota genomics, phylogenomics and proteomics

Subjects: Microbiology

Contributor: Lucia Muggia

Fungi are among the most successful eukaryotes on Earth: they have evolved strategies to survive in the most diverse environments and stressful conditions and have been selected and exploited for multiple aims by humans. The characteristic features intrinsic of Fungi have required evolutionary changes and adaptations at deep molecular levels. Omics approaches, nowadays including genomics, metagenomics, phylogenomics, transcriptomics, metabolomics, and proteomics have enormously advanced the way to understand fungal diversity at diverse taxonomic levels, under changeable conditions and in still under-investigated environments.

extremophiles fungi human opportunistic lichens

1. Introduction

The Fungi kingdom is represented by 1.5 to 5.1 million estimated species worldwide [1][2][3]. Fungi are among the most successful eukaryotes which have evolved diverse strategies even to thrive under environmental conditions where life is brought to its extremes. They have developed numerous adaptations to optimize their survival under harsh abiotic stresses [4], to colonized different substrates and to build mutualistic associations with organisms from other kingdoms (i.e., bacteria, plants and animals), thus getting advantage from the symbiotic lifestyle. Many fungi have been selected by humans for ages to be industrially exploitable organisms and are nowadays used as food or to process plant or animal materials, to produce compounds of medicinal interest or to degrade chemical compounds [5]. However, at the same time fungi can be also enemies, hardly to be defeated, as many species are serious detrimental pathogens causing economic losses to human agriculture [6], affecting animal health (human included, [7]), or damaging cultural heritages [8].

All these characteristics intrinsic of fungi require multiple changes and adaptations at deep molecular levels, which influence both the intracellular and extracellular environments. Omics approaches, nowadays including genomics, metagenomics, phylogenomics, transcriptomics, metabolomics, and proteomics have enormously advanced the way to understand fungal diversity at diverse taxonomic levels, under changeable conditions and in still under-investigated environments. These approaches can be applied both on environmental communities and on individual organism, either in nature or under in vitro conditions. In this context, cultured strains are particularly important when specific metabolic processes need to be carefully studied. However, only a minimal number of the known fungal species could be investigated for its genetic and functional diversity. Indeed, most of the taxa are difficult to retrieve in nature or even more challenging are their isolation and the stable maintenance in culture. The possibility to isolate and easily maintained certain fungal species (either yeasts or filamentous microfungi) in axenic

culture is key to facilitating thoroughly researches on their genetic and metabolic traits and have led to the selection of reference models in mycology [9].

The past decade has seen the launch of uncountable -omics projects to uncover the different aspects of fungal diversity, spanning from evolution to metabolism. Large efforts have been dedicated mainly to Saccharomycotina (being *Saccaromyces cervisiae* 'the model yeast' within Ascomycota), to several plant pathogens responsible of serious crop infections and human opportunistic species, such as *Aspergillus*, *Fusarium* or *Cryptococcus* and *Coccidioides* to mention just a few (as reviewed in [10][11]).

Due to the huge amount of studies conducted on eumycetes, this review aims at presenting an overview of the major advances in genomics, including phylogenomics, and proteomics of ascomycetes (Ascomycota), in particular reporting on examples selected from plant and animal opportunistic and pathogenic, extremophilic/polyextremotolerant filamentous and yeast-like micromycetes, as well as lichenized fungi. We also integrated this review with notions and concepts on methodological strategies and bioinformatics tools applied for sample preparations, genome and proteome sequence data analyses, respectively.

2. Towards a Genome-Based Fungal Systematics

Fungal systematics was originally based on phenotypic characters only (i.e., macro- and micro- morphology). Although most key morphological traits are fundamental for taxonomical identification, many characters may change according to abiotic growth conditions and can lead to an unreliable classification. Thanks to technological advances, the morphology-based approach has developed into an integrative taxonomic approach based on information gained from physiology, biochemistry and molecular phylogenetics, this latter based either on DNA or protein sequence data. Molecular phylogenetics has advanced enormously in the past 20 years to improving fungal systematics independently from morphology, and the application of the phylogenetic species concept (PSC) [12] lead to the recognition of uncountable new lineages at different taxonomic levels. These studies aimed at the identification of monophyletic lineages based mainly on datasets of single or multiple loci (usually up to six loci, i.e., gene trees), and tried to include both nuclear and mitochondrial markers to improve the resolution power [13][14][15][16]. However, the preferred markers have seldom considered single copy or housekeeping genes, which indeed constitute much of the cellular genome and are essential to biological functions, providing effective markers to track organismal evolution [5]. Also, phylogenetic inferences based on different loci often revealed topological incongruences, resulting in poorly supported or unresolved clades, as each gene evolves under different evolutionary pressure and time scale (e.g., [17][18]). Instead, unlinked and randomly selected orthologous loci have reconstructed robust phylogenetic hypotheses with improved accuracy [19]. Genome sequencing has therefore become essential to deliver this amount of data and in the past few years the number of available fungal genomes grew exponentially. The use of genome-wide genetic data has led to a few new proposals on how to implement species concepts. Matute and Sepulveda [20] proposed a set of standards for using genome sequences to set species boundaries, which merge identification of reciprocal monophyly, high concordance among genomic positions, lower interspecies differentiation than intraspecific diversity and low shared polymorphisms.

However, quite a long time is usually needed to gain genomic data, as genome analyses and annotation often require the settings of several parameters in bioinformatics pipelines, particularly because most fungal genomes represent still uncharted terrains. To date (6 July 2020, NCBI), approximately 6545 fungal assemblies are publicly available, of these 5230 derive from ascomycetes and were obtained from whole-genome sequences at varying degrees of completeness. Many others have been sequenced and assembled, and wait to be released (http://genome.jgi.doe.gov/programs/fungi/index.jsf). Because of this, large scale phylogenomic studies are still relatively few while more efforts have been put on genomic analyses at species and population levels [5]. Nevertheless, the development of next-generation high-throughput sequencing is greatly accelerating the access to genomic data, which become available for phylogenomic datasets and for complementing proteomic studies, to be further integrated into fungal taxonomic and systematic studies.

3. Proteomics Advances in Mycology

In the last two decades proteomics has emerged and evolved as a powerful tool for the analysis of biological systems [21]. Investigations of the proteome encompass both the identification of the protein repertoire expressed under a given physiological state in a distinct biological space at a given time, and the assessment of changes in protein abundance in response to specific sets of conditions [22][23]. Proteomics additionally involves the study of protein posttranslational modifications and protein networks. While initially focusing prevalently on the investigation of the whole-cell proteome, with advancements in the techniques, subfields of proteomics such as secretomics, subcellular, membrane and vesicle proteomics have developed and gained a crucial role in the elucidation of protein biological functions [24]. Proteomic measurements are accomplished through a combination of highly sensitive instrumentation and powerful computational methods to produce high throughput qualitative and quantitative data. A thorough work of bioinformatic data mining plays in this respect a key role: the extraction of aggregated knowledge from the data eases the way for a better understanding of the complexities of the proteome [25]. By providing information about protein levels and pathways in a given cell or a community, proteomics data have helped shedding light on organisms' eco-physiology and on the molecular basis of adaptive behaviours as well as the detection of protein biomarkers. Further aspects of the proteome, such as the dynamics of the protein components and the interactions among proteins and between proteins and other molecules, are deduced using proteomic tools which genomics and transcriptomics fail to offer [26].

In fungi, efforts toward post-genomic studies were initially made on a low number of widely investigated model organisms such as *Penicillium* sp., *Aspergillus* sp. and *Trichoderma* sp., along with their simpler relatives *Candida* and *Saccharomyces* sp. [27]. Since then, fungal proteomics research has progressed dramatically, especially due to the availability of powerful proteomics-based technologies and the advent of next-generation sequencing [28]. Much effort has also been directed towards the development of methodologies for the optimal protein extraction and separation, since fungal proteins are especially arduous to extract due to the chitin content of the fungal cell wall [29][30] Protein identification has been accomplished resorting to gel-based separation techniques coupled to mass spectrometry (MS) or tandem mass spectrometry (MS/MS) and, more recently, shotgun (gel-free) methods based on liquid chromatography (LC)-MS/MS, such as bottom-up proteomics [31][32]. These approaches have paved the

way for the development of databases collecting information about identity, relative abundances, localization and biological functions of proteins across a growing number of fungal species [33][34][35]. Fungal proteomics has consequently become an integral component of all “omics” sciences and systems biology approaches [28] to such an extent, that the quick generation of extraordinary amounts of data has outpaced the ability to assign functions. The growing disparity between known sequences and known functions for these proteins currently represents a unique challenge, where the availability of annotated genomic sequences plays a crucial role.

To date, the amount of proteomics investigations in Ascomycetes exceeds those carried out in any other fungal group. This is primarily due to the preponderance of their involvement in plant and animal diseases as well as to their multiple industrial applications [36]. Given the opportunistic and pathogenic nature of several species, proteomic analyses have been performed to further understand the biological basis of the infectious process [37] and to comprehend the mechanism required for the biologic control [38]. The biotechnological potential of fungal enzymes for the biosynthesis of products of significance has also driven an intense activity of proteomics research, more recently extended to the investigation of species from the extremes of life [39]. Several species possess excellent ability for protein production which provides one of the important aspects for identifying the protein function [40]. Furthermore, the molecular uniqueness of extremophilic and extremotolerant species has stimulated considerable interest in the search for proteins with key roles in the stress survival [41].

References

1. Blackwell, M. The fungi: 1, 2, 3 ... 5.1 million species? Am. J. Bot.2011, 98, 426–438.
2. Hawksworth, D.L. The fungal dimension of biodiversity: Magnitude, significance, and conservation. Mycol. Res.1991, 95, 641–655.
3. Hibbett, D.; Abarenkov, K.; Koljalg, U.; Opik, M.; Chai, B.; Cole, J.; et al. Sequence-based classification and identification of Fungi. Mycologia 2016, 108, 1049–1068.
4. Onofri, S.; Selbmann, L.; De Hoog, G.S.; Grube, M.; Barreca, D.; Ruisi, S.; Zucconi, L. Evolution and adaptation of fungi at boundaries of life. Adv. Space Res. 2007, 40, 1657–1664.
5. Zhang, N.; Luo, J.; Bhattacharya, D. Advances in fungal phylogenomics and their impact on fungal systematics. Adv. Genet.2017, 100, 309–328, doi:10.1016/bs.adgen.2017.09.004.
6. Dean, R.; Van Kan, J.; Pretorius, Z.A.; Hammond, K.E.; Di Pietro, A.; Spanu, P.D.; Rudd, J.J.; Dickman, M.; Khaman, R.; Ellis, J.; et al. The top 10 fungal pathogens in molecular plant pathology. Mol. Plant Pathol.2012, 13, 414–430.
7. Köhler, J.R.; Casadevall, A.; Perfect, J. The spectrum of fungi that infects humans. Cold SpringHarb. Perspect. Med.2014, 5, a019273, Cold Spring Harbor Laboratory Press, doi:10.1101/cshperspect.a019273.

8. Sterflinger, K. Fungi: Their role in deterioration of cultural heritage. Fungal Biol. Rev. 2010, 24, 47–55.
9. Perez-Nadales, E.; Nogueira, M.F.; Baldin, C.; Castanheira, S.; El Ghalid, M.; Grund, E.; Lengeler, K.; Marchegiani, E.; Mehrotra, P.V.; Moretti, M.; et al. Fungal model systems and the elucidation of pathogenicity determinants. Fungal Gen. Biol. 2014, 70, 42–67, doi:10.1016/j.fgb.2014.06.011.
10. Stajich, J.E. Fungal genomes and insights into the evolution of the kingdom. Microbiol Spectrum. 2017,5(4),FUNK-0055-2016; doi:10.1128/microbiolspec.
11. Aylward, J.; Steenkamp, E.T.; Dreyer, L.L.; Roets, F.; Wingfeld, B.D.; Wingfeld, M.J. A plant pathology perspective of fungal genome sequencing. IMA Fungus 2017, 8, 1–15, doi:10.5598/imafungus.2017.08.01.01.
12. De Queiroz, K. Species Concepts and Species Delimitation, Syst. Biol. 2007, 56, 879–886, doi:10.1080/10635150701701083.
13. Lutzoni, F.; Kauff, F.; Cox, C.J.; McLaughlin, D.; Celio, G.; Dentinger, B.; et al. (44 authors). Assembling the fungal tree of life: Progress, classification, and evolution of subcellular traits. Am. J. Bot. 2004, 91, 1446–1480.
14. Spatafora, J.W.; Sung, G.H.; Johnson, D.; Hesse, C.; O'Rourke, B.; Serdani, M.; et al. (33 authors). A five-gene phylogeny of Pezizomycotina. Mycologia 2006, 98, 1018–1028.
15. Schoch, C.L.; Sung, G.H.; Lopez-Giraldez, F.; Townsend, J.P.; Miadlikowska, J.; Hofstetter, V.; et al. (65 authors). The Ascomycota tree of life: A phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Syst. Biol. 2009, 58, 224–239.
16. Miadlikowska, J.; Kauff, F.; Högnabba, F.; Oliver, J.C.; Molnar, K.; Fraker, E. et al. (32 authors). A multigene phylogenetic synthesis for the class Lecanoromycetes (Ascomycota): 1307 fungi representing 1139 infrageneric taxa, 317 genera and 66 families. Mol. Phylogen. Evol. 2014, 79, 132–168.
17. Choi, J.J.; Kim, S.H. A genome tree of life for the Fungi kingdom. Proc. Natl. Acad. Sci. USA 2017, 114, 9391–9396, doi:10.1073/pnas.1711939114.
18. Ebersberger, I.; de Matos Simoes, R.; Kupczok, A.; Gube, M.; Kothe, E.; Voigt, K.; von Haeseler, A. A consistent phylogenetic backbone for the fungi. Mol. Biol. Evol. 2012, 29, 1319–1334, doi:10.1093/molbev/msr285.
19. Rokas, A.; Williams, B.L.; King, N.; Carroll, S.B. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 2003, 425, 798–804.
20. Matute, D.R.; Sepulveda, V.E. Fungal species boundaries in the genomic era. Fungal Gen. Biol. 2019, 131, 103249, doi:10.1016/j.fgb.2019.103249.

21. Rustagi, A.; Singh, G.; Agrawal, S.; Gupta, P.K. Proteomic studies revealing enigma of plant-pathogen interaction. In Molecular Aspects of Plant-Pathogen Interactions;Singh, A., Singh, I., Eds.; Springer: Singapore,2018; pp. 239–264, doi:10.1007/978-981-10-7371-7_11.
22. Pandey, A.; Mann, M. Proteomics to study genes and genomes. Nature 2000, 405, 837–846.
23. Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. Nature 2003, 422, 198–207, doi:10.1038/nature01511.
24. Tesei, D.; Sterflinger, K.; Marzban, G. Global Proteomics of Extremophilic Fungi: Mission Accomplished?Tiquia-Arashiro, M.G., Ed.; Springer Nature Switzerland AG 2019: Cham, Switzerland, 2019; ISBN 9783030190309.
25. Griss, J.; Perez-Riverol, Y.; Hermjakob, H.; Vizcaíno, J.A. Identifying novel biomarkers through data mining-A realistic scenario? Proteomics Clin. Appl.2015, 9, 437–443, doi:10.1002/prca.201400107.
26. Shiny, M.C.; Madhusudan, I.; Gaurav Isola, R.; Shanthi, C. Potential of proteomics to probe microbes. J. Basic Microbiol.2020, 60, 471–483, doi:10.1002/jobm.201900628.
27. Archer, D.B.; Dyer, P.S. From genomics to post-genomics in Aspergillus. Curr. Opin. Microbiol.2004, 7, 499–504.
28. Doyle, S. Fungal proteomics: From identification to function. FEMS Microbiol. Lett.2011, 321, 1–9, doi:10.1111/j.1574-6968.2011.02292.x.
29. Uranga, C.C.; Ghassemian, M.; Hernández-Martínez, R. Novel proteins from proteomic analysis of the trunk disease fungus Lasiodiplodia theobromae (Botryosphaeriaceae). Biochim. Open2017, 4, 88–98, doi:10.1016/j.biopen.2017.03.001.
30. Özhak-Baysan, B.; Ögünc, D.; Dögen, A.; Ilkit, M.; De Hoog, G.S. MALDI-TOF MS-based identification of black yeasts of the genus Exophiala. Med. Mycol.2015, 53, 347–352, doi:10.1093/mmy/myu093.
31. Bhadauria, V.; Banniza, S.; Wang, L.-X.; Wei, Y.-D.; Peng, Y.-L. Proteomic studies of phytopathogenic fungi, oomycetes and their interactions with hosts. Eur. J. Plant Pathol. 2009, 126, 81–95, doi:10.1007/s10658-009-9521-4.
32. Loginov, D.; Šebela, M. Proteomics of survival structures of fungal pathogens. N. Biotechnol.2016, 33, 655–665, doi:10.1016/j.nbt.2015.12.011.
33. Karányi, Z.; Holb, I.; Hornok, L.; Pócsi, I.; Miskei, M. FSRD: Fungal stress response database. Database 2013, 2013, bat0037, doi:10.1093/database/bat037.
34. Choi, J.; Park, J.; Kim, D.; Jung, K.; Kang, S. Fungal Secretome Database : Integrated platform for annotation of fungal secretomes. BMC Genomics2010, 11, 105, doi:10.1186/1471-2164-11-105.

35. Gudimella, R.; Nallapeta, S.; Varadwaj, P.; Suravajhala, P. Fungome: Annotating proteins implicated in fungal pathogenesis. Bioinformation2010, 5, 202–207, doi:10.6026/97320630005202.

36. Egbuta, A.M.; Mwanza, M.; Oluranti Babalola, O. A Review of the ubiquity of ascomycetes filamentous fungi in relation to their economic and medical importance. Adv. Microbiol.2016, 6, 1140–1158, doi:10.4236/aim.2016.614103.

37. Greco, T.M.; Cristea, I.M. Proteomics tracing the footsteps of infectious disease. Mol. Cell. Proteomics2017, 16, S5–S14, doi:10.1074/mcp.O116.066001.

38. Grinyer, J.; Hunt, S.; McKay, M.; Herbert, B.R.; Nevalainen, H. Proteomic response of the biological control fungus Trichoderma atroviride to growth on the cell walls of Rhizoctonia solani. Curr. Genet.2005, 47, 381–388, doi:10.1007/s00294-005-0575-3.

39. Ibrar, M.; Ullah, M.W.; Manan, S.; Farooq, U.; Rafiq, M.; Hasan, F. Fungi from the extremes of life: An untapped treasure for bioactive compounds. Appl. Microbiol. Biotechnol.2020, 104, 2777–2801, doi:10.1007/s00253-020-10399-0.

40. Sharma Ghimire, P.; Jin, C. Genetics, molecular, and proteomics advances in filamentous fungi. Curr. Microbiol.2017, 74, 1226–1236, doi:10.1007/s00284-017-1308-9.

41. Kroll, K.; Pähtz, V.; Kniemeyer, O. Elucidating the fungal stress response by proteomics. J. Proteomics2014, 97, 151–163, doi:10.1016/j.jprot.2013.06.001.