

Review Paper: Data Mining of Fungal Secondary Metabolites Using Genomics and Proteomics

Ruchi Sethi Gutch¹, Kaushal Kishore Sharma^{1*} and Aditi Tiwari²

¹Biotechnology Department, Vija Raje Government Girls P.G. College, Gwalior, (M.P.) India

²College of Veterinary Medicine, Department of Pathobiology, University of Illinois at Urbana-Champaign, USA

Abstract

Fungi are versatile organisms; they exist on earth in all extremes of conditions. Fungi are sources of important chemical entities which may be both beneficial and deleterious. Biotechnology has helped to harness this potential of Fungi in a positive direction. The advancements in Genomics and Proteomics have opened up new horizon in research. Improved advanced Molecular Biological Technologies have given a boost to our understanding of genes and helped us to exploit the full potential of Fungi. Bioinformatics and Statistical sciences are indispensable in this regard. Databases are available, providing fast, efficient, meaningful interpretation and analysis of vast amounts of data generated in scientific laboratories.

Introduction

The advancements in genomics and proteomics have opened up vast avenues to explore secondary metabolites of Fungi. The trend to sequence the entire genome of commercially important fungi in recent years has tremendous potential. Filamentous fungi, especially, are known by their ability to produce bioactive secondary metabolites. These bioactive secondary metabolites have varied applications for e.g. in brewing, baking, pharmaceutical and agro industry. The knowledge of the genes involved of a particular secondary metabolite has immense benefits in its under-expression or over-expression. The most intriguing aspect is that many fungi have clusters of genes which are not expressed. The advent of molecular biology techniques has given us the power to analyze DNA and protein sequences with greater dexterity. The products of genes which remain un-expressed could be commercially important in terms of its pharmacological and/or clinical value. [1-4]. Examination of large pre-existing databases can be done to generate new information.

Fungi an Important Source of Secondary Metabolites

Important secondary metabolites of fungi include food processing enzymes metabolites like antibiotics, organic acids and other metabolites [5]. It is crucial to decipher the entire proteins in the fungal cell system and even those that are under-expressed or unexpressed. Coupling protein separation with high resolution mass spectrometry makes it possible to have an insight about almost the entire proteins of the fungi. The incidences of antibiotic resistance are frequently encountered these days. There is an urgent need to find new therapeutic agents to combat rapidly emerging multiple drug resistance. Filamentous fungi are prolific secondary metabolite producers and have been a rich source of lead compounds for the pharmaceutical industry [6]. The harmful effects of Fungi include outbreaks of Mycotoxicoses. Mycotoxins being low weight secondary metabolites of fungi can cause great havoc. Hence, the focus of current research should be aimed at lessening the effects of mycotoxins from food and feed as their complete eradication is not possible. The aim should be to design such natural biocontrol competitive fungi and to enhance host resistance [7]. In fact, we still practically haven't realized the full biosynthetic potential of fungi. Fungal secondary metabolites can be beneficial and harmful as well, however these need to be modulated with great care and precision.

Penicillium sp.

Penicillium sp. is an important fungus producing important secondary metabolites. The morphological features of a certain strain

of *Penicillium* isolated from Brazilian cerrado soil were characterized and were distinctly different from all well-defined *Penicillium* species. The new secondary metabolites were seen to be quite efficient against *Candida albicans*, *Listeria monocitogens* and *Bacillus cereus* [8]. *Penicillium* sp. is a major source of novel secondary metabolites which are being explored using proteomics [9,10]. *Penicillium* strains can be improved via genome sequencing. DNA microarrays are used to compare transcriptomes of the sequenced strain. Many key cellular transport processes involving Penicillin and intermediates still remain too characterized at molecular level. Genes were transcriptionally up regulated under conditions that stimulate Penicillin G production, this is futuristic genomics driven metabolic engineering [11]. *Penicillium* sp. is a unique fungus possessing antibacterial activity. Earlier strain improvement was done using mutagenesis and screening, but now a days technological knowhow has expanded new horizons. Genome sequencing data along with DNA microarray technology has helped in selecting specific genes and deleting unnecessary ones [12].

Transcriptional and bioinformatics analysis of the DNA region containing penicillin gene cluster in *Penicillium chrysogenum* revealed eight transcripts which are expressed in penicillin producing conditions [13]. Industrial penicillin production can be scaled up. Crucial open reading frames have been detected, introduction of mutations was done and selecting the best recombinant strain was done to enhance production [14]. Entire genome sequencing of *Penicillium chrysogenum* (high penicillin producing) industrial strain was a very crucial development. The genome sequence has been elucidated. This has opened up new frontiers for metabolic engineering in near future [15].

Aspergillus sp.

Aspergillus, which is an important filamentous fungus, is another

*Corresponding author: Kaushal Kishore Sharma, Guest Faculty, Biotechnology Department, Vija Raje Government Girls P.G. College, Gwalior, (M.P.) India, Tel: +91-7772803265; E-mail: send2kaushal@gmail.com

Received July 31, 2015; Accepted August 29, 2015; Published September 04, 2015

Citation: Gutch RS, Sharma KK, Tiwari A (2015) Review Paper: Data Mining of Fungal Secondary Metabolites Using Genomics and Proteomics. J Data Mining Genomics Proteomics 6: 178. doi:10.4172/2153-0602.1000178

Copyright: © 2015 Gutch RS, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

significant contributor to a multibillion food industry. *Aspergillus niger* has been used to meet world's demand of citric acid. *A. oryzae* is used in both in food production (e.g. soy sauce, alcoholic beverages) and industrial enzyme production (e.g. α -amylase). *A. terreus* has been used to produce cholesterol lowering statins [16]. The 'omics' techniques available are very helpful. In the Aspergilli research post genomic studies have just begun to take place. In the proteomics approach the problem with majority of filamentous fungi is their exceptionally strong cell walls [17]. However, the ability of saprophytic fungi to thrive on diverse substrates makes them unique. These fungi are capable of producing a broad array of hydrolytic enzymes, which are mostly secreted. These enzymes have been commercially produced and sold [18,19]. Many recombinant proteins can be easily produced by these fungi as the downstream processing is less cumbersome. There is great interest in the secretome studies of *Aspergillus*. *Aspergillus* secretome is strongly dependent on both the culture conditions and nutrient sources and is a matter of great interest.

It is very intriguing that fungi possess the unique ability to digest complex molecules and mineralize various xenobiotics. Fungi have a unique approach to sense their environment and regulate their intra and extra cellular proteins, leading to diverse changes in their proteome and secretome. This unique ability to adapt itself in different harsh conditions can be closely monitored. Ascomycetes and basidiomycetes fungi, both are capable of secreting large amounts of lignocelluloses degrading enzymes; they are extremely useful in the biofuel industry [19]. Apart from food and pharmaceutical drugs such as antibiotics (namely penicillin, cephalosporine and griseofulvin) many immunosuppressant (cyclosporine A and gliotoxin) are also produced from these fungi [20] The Lae A factor, for instance from the *Aspergillus nidulans* genome, which is an important anti-tumor compound [1]. White rot fungi utilize their high redox potential to oxidize an array of environmental pollutants such as pesticides, organo chlorines, poly chlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), synthetic dyes, wood preservatives and synthetic polymers [21].

Genomics and Proteomics Analysis

At present hundreds of fungal genomes are freely available. It is the availability of sequence and genomes and the advancements in mass spectrometry (MS) that has given impetus to the fungal proteome analysis. Proteomics is a high throughput field, which relies on advanced technologies (Figure 1). Proteomics have allowed faster analysis of innumerable proteins. Protein arrays are available nowadays, which allows quantification of individual proteins in a mixture of proteins. However, in spite of fast speed, accuracy and automation, there are several limitations. In the case of oligonucleotide arrays, which have to be validated by quantitative techniques e.g. PCR (q-PCR) the protein systems also require validation, techniques such as ELISA. One more important limitation is the cost factor and availability of specific reagents for this technique. Specifically for filamentous fungi there is paucity or virtually non-availability of such antibodies for detection. Practically speaking, it is virtually impossible to distinguish protein isoforms, mutated proteins or post translationally modified proteins using this system. Comparative proteomics have been in trend nowadays to compare proteins of mutants, distantly related or unrelated fungi.

Proteomics research has progressed towards organelle proteomics. The normal procedure includes homogenization and sub-cellular fractionation. The focus has shifted towards individual organelle. The secretome studies show the wealth of enzymes secreted by filamentous

fungi. The study of secretome has greatly enhanced our knowledge of the mechanisms of metabolic processes that are significant in biotechnology. Organelle proteomics have helped in gaining an insight about compartmentalization that exists in a cell e.g. antibiotics, organic acids and protein secretion. Proteins can be quantified and different organelles can be separated. There are various constraints like a rigid cell wall in filamentous fungi. There are other problems like rigorous protein quantification, then reproducibility and performance of different steps like disruption of cells, sub-cellular fractionation are certain steps which are not done with care [22].

Proteogenomics is the area of research at the interface of Proteomics and Genomics. Designing of customized protein sequences is done via genomic and transcriptomic data. Identification of novel proteins is done which are not present in any protein sequence databases. The data generated using proteogenomics has increased tremendously (Figure 2). Although it is intriguing about how many new alternative splice forms are translated into stable functional proteins. As many protein transcripts are degraded immediately following translation. However, quantitative analysis of transcripts and protein expression data has been successfully done using Proteogenomics. No doubt this field will change and enhance our present understanding of Proteomics in near future [23].

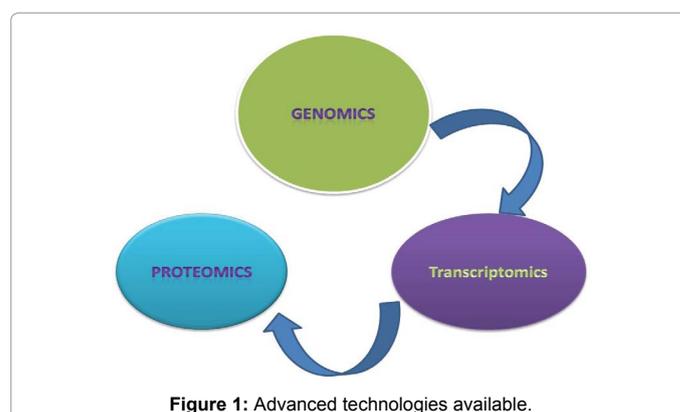
Bioinformatics and Statistics

The synergy between Bioinformatics and Statistics are often ignored and results in profound negative effect on the results. The designing of an experiment should start with clear objectives. Statistical tools should be utilized from the very start; other steps include hypothesis generation, setting up of experiment, generation of results, etc. Bioinformatics being an interdisciplinary field will help us in understanding this complex data. As processing of the biological data in the correct manner is a must. Bioinformatics combines computer science, statistics, mathematics and engineering technology to study and process biological data. Extraction of hidden predictive information from large databases by usage of Bioinformatics is being performed. Relevant data searching, analysis and interpretation can be performed to generate accurate and meaningful results. These are just a few areas where we need to concentrate neither field can efficiently function independently without others, efficient interaction is a must.

The interrelationship between different fields has been depicted in (Figure 3).

Conclusion

Fungi are eukaryotic cell factories of great importance. Genomics, Proteomics and Bioinformatics are important fields undergoing



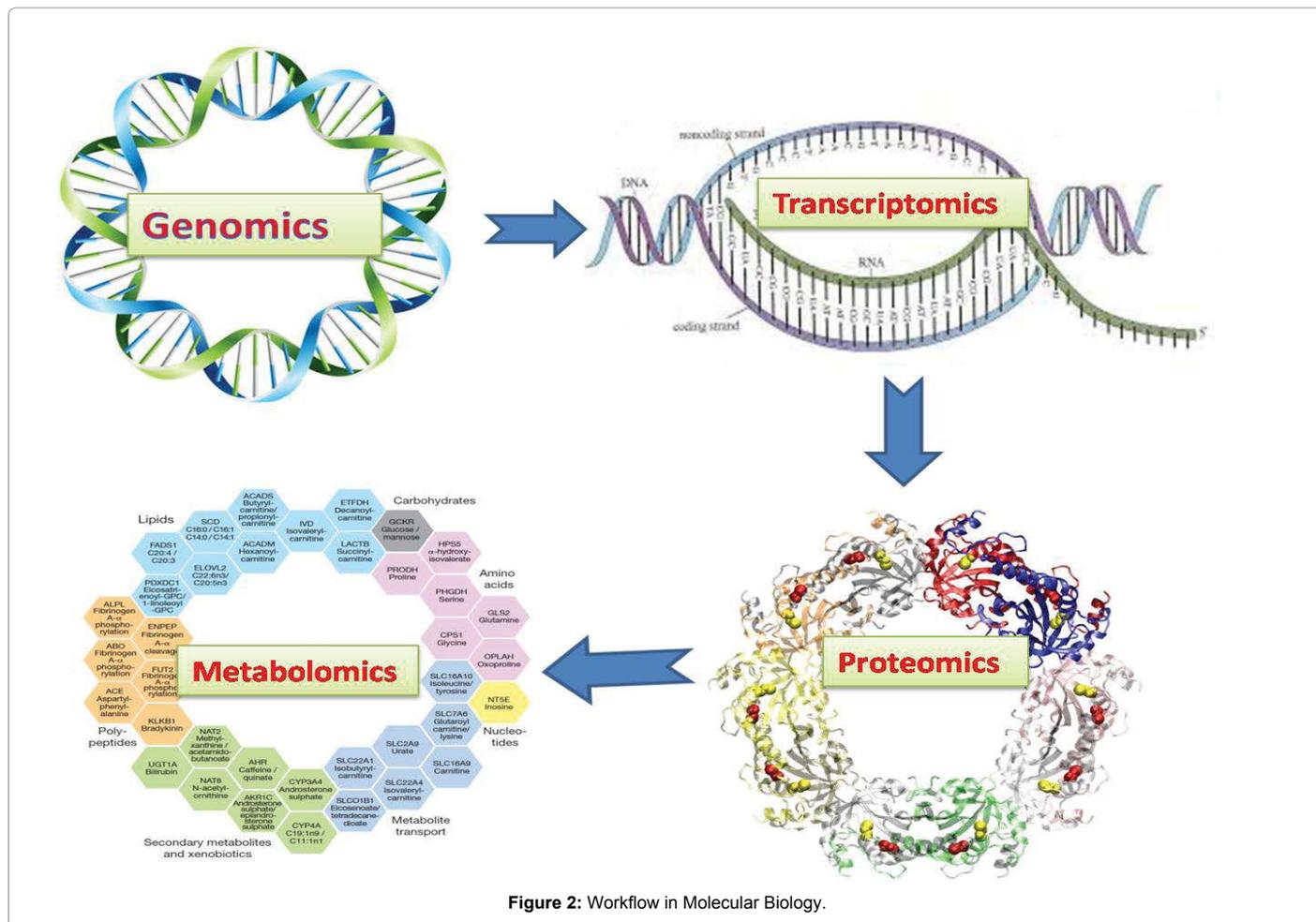


Figure 2: Workflow in Molecular Biology.

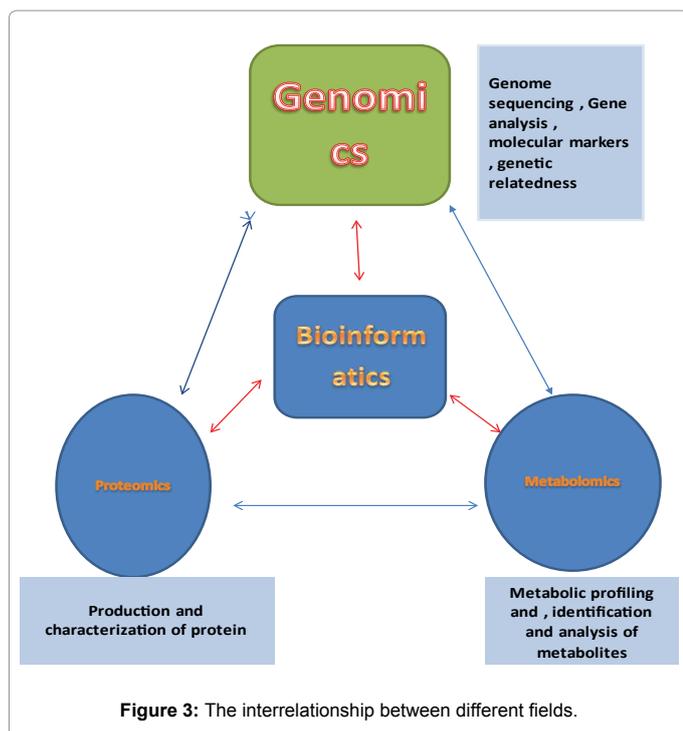


Figure 3: The interrelationship between different fields.

immense transition. This is an era of Proteogenomics where we are using customized protein sequence databases to identify novel peptides of Fungi. Bioinformatics has given us the power to handle bulk biological data. However, we need to look beyond this to harness the full potential fungi.

References

1. Bok JW, Hoffmeister D, Maggio-Hall LA, Murillo R, Glasner JD, et.al. (2006) Genomic Mining for Aspergillus Natural Products. Chem and Biol 13: 31–37.
2. Lorenzen K, Anke T (1998) Basidiomycetes as a source for new bioactive natural products. Curr Org Chem 2: 329–364.
3. Anke T (1989) Bioactive metabolites from microorganisms. In Progress in Industrial Microbiology. M.E. Bushell and U. Grafe (eds) Amsterdam: Elsevier 27: 51–66.
4. Lorenzen K, Anke T (1998) Basidiomycetes as a source for new bioactive natural products. Curr Org Chem 2: 329–364.
5. Chen XZ, Shen W, Fan Y, Wang ZX (2011) Genomics and metabolic engineering of filamentous fungi in the post-genomics era. Yi Chuan 33:1067-1078.
6. Wu MC, Law B, Wilkinson B, Micklefield J (2012) Bioengineering natural product biosynthetic pathways for therapeutic applications. Curr Opin Biotech 23: 931–940.
7. Bhatnagar D, Yu J, Ehrlich KC (2002) Toxins of filamentous fungi. Chem Immunol 81: 167–206.
8. Petit P, Lucas EMF, Ambreu LM, Pfenning LH, Takahashi JA (2009) Novel antimicrobial secondary metabolites from a *Penicillium* sp. isolated from Brazilian cerrado soil. Elect Jour of Biotech 12: 9.

9. Kobayashi A, Hino T, Yata S, Itoh TJ, Sato H, et.al. (1988) Unique spindle poisons, curvularin and its derivatives, isolated from *Penicillium* species. *Agric Biol Chem* 52: 3119–3123.
10. Kuno F, Otoguro K, Shiomi K, Iwai Y, Omura S (1996) Novel and selective acetylcholinesterase inhibitors from *Penicillium* sp. FO 4259. I. Screening, taxonomy, fermentation, isolation and biological activity. *J Antibiot* 49: 742–747.
11. van den Berg MA, Albang R, Albermann K, Badger JH, Daran JM, et.al. (2008) Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. *Nat Biotechnol* 26: 1161–1168.
12. Harris DM, van der Krogt ZA, Klaassen P, Raamsdonk LM, Hage S, et al. (2009) Exploring and dissecting genome-wide gene expression responses of *Penicillium chrysogenum* to phenylacetic acid consumption and penicillinG production. *BMC Genom* 10: 75.
13. Fierro F, García-Estrada C, Castillo NI, Rodríguez R, Velasco-Conde T, et.al. (2006) Transcriptional and bioinformatic analysis of the 56.8 kb DNA region amplified in tandem repeats containing the penicillin gene cluster in *Penicillium chrysogenum*. *Fungal Genet Biol* 43: 618–629.
14. van den Berg MA, Westerlaken I, Leeflang C, Kerkman R, Bovenberg RA (2007) Functional characterization of the penicillin biosynthetic gene cluster of *Penicillium chrysogenum* Wisconsin54-1255. *Fungal Genet Biol* 44: 830–844.
15. Wang FQ, Zhong J, Zhao Y, Xiao J, Liu J, et.al. (2014) Genome sequencing of high-penicillin producing industrial strain of *Penicillium chrysogenum*. *BMC Genom* 15: S11.
16. Kim Y, Nandakumar MP, Marten MR (2008) The state of proteome profiling in the fungal genus *Aspergillus* Briefings in Functional Genomics and Proteomics. 7: 87–94.
17. Bowman SM, Free SJ (2006) The structure and synthesis of the fungal cell wall. *Bioess* 28: 799–808.
18. Nandakumar MP, Marten MR (2002) Comparison of lysis Methods and preparation protocols for 1-D and 2-D electrophoresis of *Aspergillus oryzae* intracellular proteins. *Electrophor* 23: 2216–2222.
19. Anderson WF, Akin DE (2008) Structural and chemical properties of grass lignocelluloses related to conversion for biofuels. *J Ind Microbiol Biotechnol* 35: 355–366.
20. Bianco L, Perrotta G (2015) Methodologies and perspectives of proteomics applied to filamentous fungi: from sample preparation to secretome analysis. *Int J Mol Sci* 16: 5803–5829.
21. Mai C, Kues U, Miltz H (2004) Biotechnology in the wood industry. *Appl Microbiol Biotechnol* 63: 477–494.
22. de Oliveira JM, de Graaff LH (2011) Proteomics of industrial fungi: Trends and insights for biotechnology. *Appl Microbiol Biotechnol* 89: 225–237.
23. Nesvizhskii AI (2014) Proteogenomics: Concepts, applications, and computational strategies. *Nat Methods* 11: 1114–1125.