

Chapter 1

Proteoinformatics and Agricultural Biotechnology Research: Applications and Challenges



Jameel R. Al-Obaidi

Contents

1.1 Introduction.....	1
1.2 Proteoinformatics in Plant Disease Management.....	2
1.3 Proteoinformatic Databases and Tools.....	8
1.4 Protein-Protein Interaction Software and Database.....	10
1.5 Proteoinformatics of Edible Mushroom.....	14
1.6 Proteoinformatics of Animal Breeding Programs.....	14
1.7 Conclusion.....	15
References.....	16

1.1 Introduction

Bioinformatics is a multidisciplinary field incipient from the interaction of information, statistics, and biological sciences to analyze genome and/or proteome contents, sequence information, and predict the function and structure of cellular molecules that are used in construing genomics and proteomics information from an agricultural organism (Benton 1996; Bruhn and Jennings 2007). Bioinformatics is considered relatively new yet is a significant discipline within the biological sciences that has offered scientists and agrobiologists to interpret and handle huge amounts of information (Bartlett et al. 2016). This amount of data produced lead to the advancement and development of bioinformatics. The multi-“omics,” together with computational biology are considered important tools in understanding genomics and its products which trigger several animal, plant, and microbial functions (Mochida and Shinozaki 2011). The functional analysis for those organisms includes profiling of gene products, prediction of interaction between proteins and their sub-cellular location, and also the prediction of protein metabolism pathway simulation (Xiong 2006). Bioinformatics as a tool is not isolated but frequently interacts with

J. R. Al-Obaidi (✉)

Agro-Biotechnology Institute Malaysia (ABI), National Institutes of Biotechnology Malaysia (NIBM), 43400, Serdang, Selangor, Malaysia
e-mail: jameel@nibm.my

© Springer Nature Switzerland AG 2019

K. R. Hakeem et al. (eds.), *Essentials of Bioinformatics, Volume III*,
https://doi.org/10.1007/978-3-030-19318-8_1

other biological sections to produce assimilated results. For example, prediction of the structure of a protein depends on gene sequence and gene expression profiles, which require the use of phylogeny tools in sequence analysis. Therefore, the field of bioinformatics has developed in a way that the most important duty now include the interpretation of different types of information, including nucleotide and protein sequence and protein structures and function (Moorthie et al. 2013; Merrill et al. 2006). The analysis of DNA/RNA sequences, protein sequences and function, genome analysis and gene expression, and protein involvement in physiological functions can all make use of bioinformatic methods and tools and cannot be done without it (Collins et al. 2003). Protein sequence information and its related nucleic acid and data from many agricultural species deliver a substance for agricultural research leading to a better understanding of global agricultural needs and challenges (Kumar et al. 2015). Utilizing the available information allows and assists the identification of expression of a gene which may help to understand the relationship between phenotype and genotype (Orgogozo et al. 2015). The involvement of proteomic applications for analyses of crop, animals, and microorganisms has rapidly increased within the last decade (Mochida and Shinozaki 2010). Although proteomic approaches are regularly used in plant research worldwide, and establish powerful tools, there is still a significant area for improvement.

Proteoinformatics could be defined as “utilization of computational biology tools in the study of the proteome.” Proteoinformatics is a field involving mathematics, programing sciences, statistics, and protein biology and biochemistry to predict and analyze their structure, function, and role in cell physiology (Cristoni and Mazzuca 2011; Hamady et al. 2005). Since the data obtained from agricultural proteomic research are complex and massive in size, the role of proteoinformatics is essential to reduce the time for investigation and to deliver statistically significant results and that will help to improve the plant/animal quality based on healthy growth and high productivity. Thus, proteoinformatics is a dynamic field for the development of new breed’s diagnostic tools in order to develop pathogen-free/resistance and abiotic stress tolerance, high-quality traits, and higher quantity production (Koltai and Volpin 2003).

1.2 Proteoinformatics in Plant Disease Management

Among different plant pathogens, such as viruses, bacteria, and oomycetes, fungi are considered the most destructive (Dangl and Jones 2001). The growth, propagation, and survival strategies of pathogens are varied, but the strategies, in general, are similar, which start by colonization and progress to overcome host defense system and then finally infection establishment (Pegg 1981; Lawrence et al. 2016). As a result, the host-pathogen systems have led to a complex relationship between the host and the pathogen molecules, resulting in relationship with a high degree of variation (Hily et al. 2014). Proteomic studies focused mainly on the response of host plant upon pathogen attack that opened up a new era for biology in general and

for agriculture in particular (Lodha et al. 2013; Alexander and Cilia 2016). Along with the use of proteomic approaches in agricultural research and the progress in sequencing agriculturally important organisms, the combination of bioinformatics and proteomics generally enhance the research in this area. This kind of multidisciplinary research is likely to fill in the gap toward the understanding of host-pathogen interaction network (Koltai and Volpin 2003). Two-dimensional gel electrophoresis has been initially used for rapidly identifying major proteome differences in control versus inoculated plants. Although many proteins identified during host-pathogen interactions have been highlighted, majority are known previously and are mainly in host immunity mechanism (Memišević et al. 2013). However, those results that arise from proteomic-based research are of great significance for the validation of gene expression in genomic or transcriptomic studies (Nesvizhskii 2014). Nevertheless, by using the gel-based proteomic tools, little novel information has been obtained, especially due to the lack of sufficient bioinformatics-related information such as genome sequences (Cho 2007). Indeed, only the most abundant proteins are detected in two-dimensional gels and successfully identified by mass spectrometry (MS). Therefore, a gap seems to be in the bioinformatics channel for the proteomics research of organisms without complete genome sequencing (Sheynkman et al. 2016). These information-related limitations in agricultural proteomic research need to be overcome to increase our knowledge on protein expression during plant-microbe interactions. However, proteomic tools have grown rapidly, and new approaches and apparatus are being developed (Mehta et al. 2008; Pérez-Clemente et al. 2013). Previous agricultural proteomics research, which mainly focused on model crops, has provided fundamental understandings into different protein families in agri-organism systems' modification and regulation (Hu et al. 2015; Vanderschuren et al. 2013). Nonetheless, model crop research itself does not retain all the information and data of interest to agricultural biology (Mirzaei et al. 2016; Carpentier et al. 2008). Therefore, those crops without complete genome sequence or sufficient genomic/EST information freely available need to be investigated (Ke et al. 2015; Ekblom and Wolf 2014). In comparison to the model organisms related to agriculture, such as rice (Koller et al. 2002), maize (Pechanova et al. 2010), chicken (Burgess 2004), cattle (Assumpcao et al. 2005), brewer's yeast (Khoá Pham and Wright 2007), and the plant pathogen *Botrytis cinerea* (Fernández-Acero et al. 2009), non-model species with little or no "bioinformation" was largely affected when it comes to proteomic analysis (Armengaud et al. 2014). Economic significance and the complexity of the genome make it necessary to sequence that organism (Bolger et al. 2014), but that is not enough to make it as a model organism if that information is not reachable by the scientific community (Canovas et al. 2004), Table 1.1 shows proteomic study of non-model organism. Most mass spectrometry proteomic methods depend on complete sequence for identification; for that reason, the analysis of these non-model species remains a challenge. Thus, relying on complete and comprehensive established database for the closely related model species "conserved genome region within the species of family" will be the only choice (Hutchins 2014; Zhu et al. 2017; Bischoff et al. 2016). However, sequence variation remains an issue, especially for quantitative proteomics

Table 1.1 Proteomic studies on non-model-pathogen interaction (2008–2018)

Plant-pathogen interaction	Proteomic platform	Main findings	References
<i>Phaseolus vulgaris</i> - <i>Uromyces appendiculatus</i>	LC-MS/MS	Resistance-genes are part of the basal system and repair disabled defenses to reinstate strong resistance	Lee et al. (2009)
cacao leaves- <i>Moniliophthora perniciosa</i>	1-DE and 2-DE	Protocols described in the study could help to develop high-level proteomic and biochemical studies in cacao also being applicable to other recalcitrant plant tissues	Pirovani et al. (2008)
<i>Capsicum chinense</i> -pepper mild mottle virus (PMMoV)	2-DE and MALDI-TOF/TOF	Evidence is presented for a differential accumulation of C. chinense PR proteins and mRNAs in the compatible (PMMoV-I)-C. chinense and incompatible (PMMoV-S)-C. chinense interactions for proteins belonging to all PR proteins detected	Elvira et al. (2008)
Citrus-citrus sudden death virus (CSDaV)	2-DE and MALDI-TOF/TOF	Downregulation of chitinases and proteinase inhibitors in CSD-affected plants is relevant since chitinases are well-known pathogenesis-related proteins, and their activity against plant pathogens is largely accepted	Cantú et al. (2008)
<i>Beta vulgaris</i> -Beet necrotic yellow vein virus (BNYVV)	MALDI-TOF-MS	Involvement of systemic resistance components in Rz1-mediated resistance and phytohormones in symptom development	Larson et al. (2008)
<i>Glycine max</i> - <i>Heterodera glycines</i>	2DE and ESI/MS-MS	Differed in resistant and susceptible Soybean Roots without cyst nematode (SCN) infestation and may form the basis of a new assay for the selection of resistance to SCN in soybean	Afzal et al. (2009)
<i>Glycin max</i> - <i>Bradyrhizobium japonicum</i> and <i>Phytophthora sojae</i>	2DE and quadrupole TOF MS/MS	Sap proteins from soybean that are differentially induced in response to <i>B. japonicum</i> and <i>P. sojae</i> elicitor treatments and most them were secreted proteins	Subramanian et al. (2009)
<i>Phoenix dactylifera</i> - <i>Beauveria bassiana</i> , <i>Lecanicillium dimorphum</i> and <i>L. cf. psalliotae</i>	2DE and MALDI/TOF-TOF	Proteins related with photosynthesis and energy metabolism in date palm were affected by entomopathogenic fungi colonization	Gómez-Vidal et al. (2009)

(continued)

Table 1.1 (continued)

Plant-pathogen interaction	Proteomic platform	Main findings	References
<i>Solanum lycopersicum</i> —Cucumber Mosaic Virus (CMV)	DIGE and MS/MS analysis	The study demonstrated that virus infection in transgenic tomato is restricted to the inoculated leaves. The study contributes to defining the role played by key proteins involved in plant-virus interaction and to studying antibody-mediated resistance	Di Carli et al. (2010)
<i>Saccharum officinarum</i> - <i>Gluconacetobacter diazotrophicus</i>	SDS-PAGE and ESI-Q-TOF	30 identified bacterial proteins in the roots of the plant samples; from those, 9 were specifically induced by plant signals	Lery et al. (2010)
<i>Brassica juncea</i> - <i>Albugo candida</i>	2DE and Q-TOF MS/MS	The study demonstrates that the timing of the expression of defense-related genes plays a crucial role during pathogenesis and incompatible interactions and that the redox balance within the chloroplast may be of crucial importance for mounting a successful defense response	Kaur et al. (2011)
<i>Citrus aurantifolia</i> - <i>Candidatus Phytoplasma aurantifolia</i>	2-DE and MS	The study provided proteomic view of the molecular basis of the infection process and identify genes that could help inhibit the effects of the pathogen	Taheri et al. (2011)
<i>Gossypium barbadense</i> - <i>Verticillium dahliae</i>	2-DE, EST database-assisted PMF and MS/MS	Infection causes elevation in ethylene biosynthesis, Bet v 1 family proteins may play an important role in the defense reaction against <i>Verticillium</i> wilt, and wilt resistance may implicate the redirection of carbohydrate flux from glycolysis to pentose phosphate pathway (PPP)	Wang et al. (2011)
<i>Vitis vinifera</i> - <i>Uncinula necator</i>	2-D DIGE And	High levels of Mn concentration in grapevine leaves triggered protective mechanisms against pathogens in grapevine	Yao et al. (2012)
<i>Brassica napus</i> - <i>Sclerotinia sclerotiorum</i>	2-DE and MALDI TOF/TOF	The study showed new insights into the resistance mechanisms within <i>B. napus</i> against <i>S. sclerotiorum</i>	Garg et al. (2013)
<i>Citrus unshiu</i> - <i>Penicillium italicum</i>	2-DE and LC-QToF-MS	Lignin plays important roles in heat treatment-induced citrus fruit resistance to pathogens	Yun et al. (2013)
<i>Solanum lycopersicum</i> - <i>Pseudomonas syringae</i>	iTRAQ	The study provided an insight into tomato's response to <i>Pseudomonas syringae</i>	Parker et al. (2013a)

(continued)

Table 1.1 (continued)

Plant-pathogen interaction	Proteomic platform	Main findings	References
<i>Mentha spicata</i> - <i>Alternaria alternata</i>	2-DE and MALDI TOF-TOF MSMS	The study deciphers the mechanism by which a foreign metabolite mediates stress tolerance in plant under control and infected condition	Sinha et al. (2013)
Beta vulgaris- Beet necrotic yellow vein virus	LC-MS/MS	The study identified proteins associated with systemic acquired resistance and general plant defense response	Webb et al. (2014)
<i>Phytophthora infestans</i> - <i>Solanum tuberosum</i>	LC-MS/MS	Proteins involved in sterol biosynthesis were downregulated, whereas several enzymes involved in the sesquiterpene phytoalexin biosynthesis were upregulated	Bengtsson et al. (2014)
<i>Anacardium occidentale</i> - <i>Lasiodiplodia theobromae</i>	2DE- SI-Q-TOF MS/MS	Cashew responsive proteins indicate modulation of various cellular functions involved in metabolism, stress/defense, and cell signaling	Cipriano et al. (2015)
<i>Lactuca sativa</i> - <i>Salmonella enterica</i>	2DE and nano LC-MS/MS	Proteins involved in lettuce's defense response to bacterium were upregulated, such as pyruvate dehydrogenase, 2-cys peroxiredoxin, and ferredoxin-NADP reductase	Zhang et al. (2014b)
<i>Oil palm</i> - <i>Ganoderma interaction</i>	2DE, MALDI TOF/TOF	Proteins related to lignin synthesis were downregulated up on interaction	Al-Obaidi et al. (2014)
<i>Amorpha fruticosa</i> - <i>Glomus mosseae</i>	iTRAQ and LC-MS/MS	77 proteins were classified according to different functions during the interaction	Song et al. (2015)
<i>Vitis vinifera</i> - <i>Xylella fastidiosa</i>	2DE, MALDI TOF/TOF	Muscadine and Florida hybrid bunch grapes express novel proteins in xylem to overcome pathogen attack	Katam et al. (2015b)
<i>Solanum tuberosum</i> - <i>Ralstonia solanacearum</i>	2DE and MALDI-TOF/TOF	The study showed involvement of the identified proteins in the bacterial stress tolerance in potato	Park et al. (2016)
<i>Malus domestica</i> - <i>Botryosphaeria berengeriana</i>	2DE and MALDI-TOF-TOF	The study speculated that the upregulation of abscisic stress ripening-like protein and the dramatic decrease of -adenosylmethionine synthetase in the resistant host could be related to its better disease resistance	Cai-xia et al. (2017)
<i>Paulownia fortunei</i> - <i>Phytoplasma</i>	iTRAQ	Paulownia witches' broom (PaWB) proteins may help in developing a deeper understanding of how PaWB affects the morphological characteristics of <i>P. fortunei</i>	Wei et al. (2017)

approaches, which will lead to low coverage of protein identification (Chandramouli and Qian 2009; Zhan et al. 2017). Moreover, “conserved genome” regions may produce similar protein sequence with different cellular functions and may increase the number of mismatch protein identities (Khan et al. 2014). Gel-based proteomics is considered the most dominant platform used for agricultural proteomic research (Tan et al. 2017). However, the use of gel-free proteome analysis is increasing rapidly in agricultural research with the presence of more proteoinformatics data (Porteus et al. 2011; Komatsu et al. 2013). Pathogen proteins that are used to suppress host defenses are of high importance in agricultural host-pathogen interaction, as these proteins may play a role in virulence, pathogenicity, and effector molecules (Van De Wouw and Howlett 2011). Pathogen characteristics are of primary interest in crop development programs (Fletcher et al. 2006). The contribution of proteoinformatic advances has helped the sequencing of the entire genomes of many pathogens in the last 10 years (Land et al. 2015). Classical biochemistry and molecular biology, as well as the modern omic platform techniques coupled with bioinformatic tools research, have been conducted on agricultural-related pathogens and their interactions with crops (Barah and Bones 2014). Recently, the study of pathogens have been significantly promoted by the availability of bioinformatic data and the resources for multi-“omics” research (Bhadauria 2016). These approaches, in combination with gene-targeting studies such as targeted mutations and gene silencing studies, are explained in molecular host-pathogen communications and the complex mechanisms involving pathogenesis and virulence (Allahverdiyeva et al. 2015; McGarvey et al. 2009; Fondi and Liò 2015). The present efforts to provide sufficient “proteoinformation” to determine related proteins and their function have improved the capacity to understand the core causes of crop and animal diseases and develop new possibilities of treatments (Chen et al. 2010). Proteoinformatics has many practical applications in current agricultural-related disease management with respect to the study of host-pathogen interactions, understanding the nature of the disease genetics, pathogenicity, and/or virulence factor of a pathogen which eventually aid in designing better disease control and drive the infection process which has also been identified, using molecular biological technologies and genetics in identifying the interaction with bacteria such as tomato and *Pseudomonas syringae* (Parker et al. 2013b; Balmant et al. 2015) and rice and *Xanthomonas oryzae* (Wang et al. 2013b) or with virus such as potato and potato virus (PVY) (Stare et al. 2017) or with phytopathogenic fungi such as apple and *Alternaria alternate* (Zhang et al. 2015), strawberry and *Fusarium oxysporum* (Fang et al. 2013), cotton and rot fungus *Thielaviopsis basicola* (Coumans et al. 2009), and coffee and *Hemileia vastatrix*. Proteoinformatic tools and databases related to agricultural diseases need to be further developed and expanded. Obviously, tools, software, and databases are adapted from human and more specifically medical analysis systems, and these may not necessarily be a model for analysis of crop proteomic data; therefore, more information regarding those crops and their pathogens will be very helpful to fill in the proteoinformation gap in agricultural research and also to verify the protein information predicted in the literature (Dennis et al. 2008; Thrall et al. 2011; Van Emon 2016). Generally, the proteoinformation is larger and more complicated

than the genoinformation, especially in crops, since there are more proteins than genes. That is mainly because of the post-translational enzymatic modification. The nucleotide sequence can represent the genome of an organism; on the other side, peptide sequence cannot represent the proteome for that organism unless the structure of an interaction between those proteins revealed (Gupta et al. 2007; Khan 2015).

1.3 Proteoinformatic Databases and Tools

Sequencing projects of crops and animals related to agriculture bring the number of proteomic research in this field higher. Proteoinformatic methods and tools could be used to identify a specific protein of interest within the proteome of an organism which could be valuable for community related to agriculture and to interpret their cellular functions. The different and unusual protein information might be used to develop drought- and salt-tolerant crops, for diseases resistance and improvement of livestock, and higher productivity (Fears 2007; Gong et al. 2015; Ahmad et al. 2016). As discussed, a closely related sequence for a specific crop or animal can be used if genome information is not accessible. The ever-growing databases of whole genome sequence remain to accelerate capabilities of proteoinformatics, till the time of writing this chapter; there are more than 500 plants with whole genome sequence from more than 5000 eukaryotic sequence since the first genome sequence of plant (*Arabidopsis*) in the year 2000 (Kaul et al. 2000). Bioinformatic investigations of the genome-based information from important commercial crops revealed that gene organization over evolutionary time has remained constant and conserved, which means that knowledge obtained from model plants such as *Oryza sativa* and *Arabidopsis thaliana* may be exploited to propose food improvement programs for monocot and dicot crops, respectively (Ong et al. 2016; Jayaswal et al. 2017).

In proteoinformatics, the term “peptide/protein sequence” implies subjecting those sequences or its related databases or other methods of bioinformatics on a computer. Sequence alignment in proteoinformatics is ordering the sequences of protein/peptide, RNA, or DNA to find similar regions that may be a sign of functional and structural relationship (Pearson 2013), some important proteoinformatics databases listed in Table 1.2.

Proteoinformatics is considered as an evolving field of agricultural research. Interpreting particular functions of crops/animals is essential to determine useful proteins to improve agricultural traits (Newell-McGloughlin 2008). The integration of proteoinformatics and other omic platforms databases from agricultural species is of high importance to promote/enhance crops/animals system to solve global issues such as food, water stress, and climate changes (Katam et al. 2015a). For Asia, for instance, the Asia Pacific Bioinformatics Network (www.apbionet.org) is a good regional source (Khan et al. 2013).

Besides the classical well-known database, many website-based database or platform content have served proteomics and have been used in agricultural research.

Table 1.2 Proteoinformatics online databases/resources

Database/resources	Website link	References
Protein Information Resource	http://pir.georgetown.edu/	Wu et al. (2003)
Protein Knowledgebase	www.uniprot.org/	The UniProt Consortium (2017)
Protein domain database	http://prosite.expasy.org/	Hulo et al. (2008)
Database of Interacting Proteins	http://dip.doe-mbi.ucla.edu/dip/Main.cgi	Salwinski et al. (2004)
Large collection of protein families	http://pfam.xfam.org/	Finn et al. (2016)
Protein fingerprints	http://130.88.97.239/PRINTS/index.php	Attwood et al. (2012)
Protein data bank	http://www.wwpdb.org/	Gore et al. (2017)
Server and Repository for Protein Structure Models	https://swissmodel.expasy.org/	Biasini et al. (2014)
Database of Comparative Protein Structure Models	https://modbase.compbio.ucsf.edu/modbase-cgi/index.cgi	Pieper et al. (2014)
A General Repository for Interaction Datasets	https://thebiogrid.org/	Chatr-aryamontri et al. (2017)
Comprehensive Enzyme Information System	http://www.brenda-enzymes.org/	Placzek et al. (2017)
Encyclopedia of Genes and Genomes	http://www.kegg.jp/	Kanehisa et al. (2012)
Interacting Genes/Proteins	https://string-db.org/	Szklarczyk et al. (2017a)

The ExPASy Proteomics site, for instance, is considered as a tool developed for human proteomic research (Gasteiger et al. 2003; Hoogland et al. 2007); however, it is widely used to compute isoelectric point (pI) and molecular weight (Mw) for agricultural proteomic studies (Imam et al. 2014; Dahal et al. 2010; Guijun et al. 2006; Schneider et al. 2004; Lande et al. 2017). In general, regarding agricultural proteomics, there are a number of web-based proteomics databases that hold a plenty of efficient information (Martens 2011). Recently, a new website was developed for tracking information and articles related to the changes in plant proteomes in response to stress (PlantPReS; www.proteome.ir). Organelle proteomic analyses have also been performed in animal and plant databases such as Organelle DB (<http://labs.mcdb.lsa.umich.edu/organelledb/>) (Agrawal et al. 2011). Organelle expression proteomics was considered as successful tools focusing on subcellular proteins rather than total proteins (Yates Iii et al. 2005) such as mitochondrial proteome research in potato (Salvato et al. 2014), chloroplast in tomato (Tamburino et al. 2017), endoplasmic reticulum in rice (Qian et al. 2015), peroxisomes in spinach (Babujee et al. 2010), vacuoles in cauliflower (Schmidt et al. 2007), and nucleus in soybean (Cooper et al. 2011) because they have fewer proteins which can easily be identified since they contain a limited number of proteins; thus, protein identification will be more appropriate. In the last 30 years, gel-based proteomics has been used as a main platform for agricultural proteomics. The gel is stained to visualize

the proteins that have travelled to specific locations in the gel. For complex samples, proteins are analyzed after enzymatic digestion (Padula et al. 2017). Many software programs were developed for gel analysis (single stained and 2D-DIGE) and used in many agricultural-related proteomic research, most of which are commercial software such as Delta2D (<http://www.decodon.com/delta2d.html>), ImageMaster 2D Platinum, Melanie 9 (<http://2d-gel-analysis.com/>), PDQuest (<http://www.biorad.com/en-ch/product/pdquest-2-d-analysis-software>), Samsspots, SpotsQuest and SpotMap (<http://www.cleaverscientific.com>), and Dymension (<http://www.syn-gene.com/dymension>). While some of the free available software have not survived and they are either not available for download or totally discontinued such as Gel IQ from (<http://ludesi.com/>), there are few software which are still available and functioning (Maurer 2016; Singh 2015) such as Gel2DE, SDA for DIGE analysis, and RegStatGel (<http://www.mediafire.com/FengLi/2DGelsoftware>).

Followed by protein separation, the peptide MS/MS fragmented spectra are matched against the available sequence in the database for protein identification. The peptide sequence identification is obtained based on the similarity score among the experimental MS/MS and the theoretical MS/MS spectra. The mass spectra obtained during protein identification are matched with the hypothetical one existing in the database and a statistical score, based on the spectrum resemblance, is associated with the protein identification. The restraint of this approach is that only known proteins/genes reported in the database can be identified (Nilsson et al. 2010). Recently, NCBI dropped “gi number” identifier and replaced the NCBI nr database with a newer database named NCBIProt which is more complicated yet more comprehensive (Disruption ahead for NCBI databases 2016). The only disadvantage of this new database is that it is time-consuming to search for non-model organism although slight improvement was noticed (data not shown). De novo sequencing can be the method of choice when the protein, in this case, the sequence is obtained directly from the MS/MS spectra to skip the step of database spectrum search. The resulted sequences are then compared with those contained in the database so to detect homologies (Ekblom and Wolf 2014).

Database search software programs/tools is listed in Table 1.3 together with those employed for de novo searching. An example of software used for de novo peptide sequencing is the Novor (www.rapidnovor.org/novor), which is capable of performing real-time de novo sequence analysis with high accuracy (Ma 2015).

1.4 Protein-Protein Interaction Software and Database

Physiological and molecular cell processes are mainly carried out through the interactions between different proteins. Interactions are physical relations between different protein structures via weak bonds (Khazanov and Carlson 2013; Chang et al. 2016). In agricultural proteomic research, identifying protein identities binding or interacting with each other during certain defined circumstances and determining

Table 1.3 List of mass spectrometry search-related software/websites

Software	Description	Website/download link	References
Maxquant	A quantitative proteomics aimed at high-resolution MS data	http://www.coxdocs.org/doku.php?id=maxquant:common:download_and_installation	Cox et al. (2011)
Byonic™	Full MS/MS search engine providing comprehensive peptide and protein identification	https://www.proteinmetrics.com/products/byonic/	Bern et al. (2002)
Mascot	A platform able to read various binary mass spectrometry data files	http://www.matrixscience.com/	Cottrell (2005)
MassMatrix	Database search algorithm for tandem mass spectrometric data	www.massmatrix.bio	Xu and Freitas (2009)
MS Amanda & Elutator	Scoring system to identify peptides out of tandem mass spectrometry	http://ms.imp.ac.at/i	Doblas et al. (2017)
Cyclobranch	De novo engine for identification of nonribosomal peptides	http://ms.biomed.cas.cz/cyclobranch/docs/html/	Novák et al. (2015)
Maxquant	Quantitative proteomics software for analysis of label-free and SILAC-based proteomics	https://web.archive.org	Tyanova et al. (2016a)
SWATH	Commercial software processing tool within PeakView data can be exported for statistical analysis after false discovery rate analysis	https://sciex.com/technology/swath-acquisition	Kang et al. (2017)

the protein-binding site are of very high importance for a better understanding of the bases of many biological/physiological activities.

Protein interactions play a significant role in protein characterization and the discovery of protein functions and the pathways they are involved in (Rao et al. 2014). This is especially true during mutualism (symbiotism), commensalism, and parasitism interaction which is caused by specific protein-protein interactions (PPI) between organisms (Leung and Poulin 2008). The precision of experimental results in revealing protein-protein interactions, however, is rather doubtful, and the availability of high-throughput platforms has shown inaccuracy and false-positive information for protein interaction. Considering experimental restrictions and limitation to find all interactions in a specific proteome, computational prediction of protein

interactions is a requirement to proceed on the way to complete interactions at the proteome level (Keskin et al. 2016). Affordability of high-throughput machines and the development of computational-based prediction methods have produced vast numbers of protein-protein interactions. Computational methods for protein-protein interaction predictions can use a variety of biological data gene and protein sequences, evolution, and expression. Algorithms and statistics are commonly used to assimilate these data and deduce PPI predictions (Clark et al. 2011). This ability to provide comprehensive and reliable sets of PPIs prompted the development of many databases, aiming to gather and unify the available data, each with a different focus and different strengths. List of PPI database and examples in agriculture are presented in Table 1.4. Protein-protein interaction has been investigated and studied in many agricultural-related research such as rice with specific network (<http://bis.zju.edu.cn/prin/>)(Gu et al. 2011; Zhu et al. 2011), *Rhizoctonia solani*-rice interaction (Lei et al. 2014), maize (<http://comp-sysbio.org/ppim/>) (Zhu et al. 2015), chicken, and cattle (Fen et al. 2016).

One of the most common databases in agricultural research is the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (Szklarczyk et al. 2015, 2017b); it is another database that incorporates both known and predicted network between proteins. Currently, STRING database covers more than 2000 species, and it is expected to cover more than 4000 in its 11th version (current version 10.5). STRING can give 3D structure besides the interaction network of a given proteome, the database used widely in prediction of protein interaction in agricultural proteomic-related research such as crop under biotic stress (Liu et al. 2015; Vu et al. 2016; Al-Obaidi et al. 2016a; Wu et al. 2015), oil-crop metabolism (Raboanatahiry et al. 2017), phytopathogenic fungi (Chu et al. 2016; Li et al. 2017), mushroom cultivation (Rahmad et al. 2014), poultry (Broiler chicken) (Zheng et al. 2016;

Table 1.4 List of protein-protein interaction (PPI) software/website

STRING	Provide a critical assessment and integration of protein-protein interactions	https://string-db.org/	Szklarczyk et al. (2017b)
MENTHA	Provides protein-protein interaction (PPI) data for many species	http://mentha.uniroma2.it/	Calderone and Cesareni (2012)
GPS-Prot	Computational prediction of phosphorylation sites	http://gps.biocuckoo.org/	Xue et al. (2008)
Compass	This tool is applicable to proteomic investigations ranging from focused studies on a small number of selected proteins	http://www.proteinsimple.com/compass/downloads/	Wenger et al. (2011)
Perseus	Shotgun proteomics data analyses	http://www.coxdocs.org/doku.php?id=perseus:start	Tyanova et al. (2016b)
Struct2Net	Structure-based protein-protein extraction	http://cb.csail.mit.edu/cb/struct2net/webserver/	Hosur et al. (2012)

Zheng et al. 2014), and buffalo (Ashok and Aparna 2017). The interactive STRING network can be recalculated based on user setting and cut-off values as well as interaction score, the maximum number of shown interactions, and expanded based on user selected. Currently, it is not clear whether protein-protein interaction networks and database are representing the true biological interactomes. For that reason, agricultural proteomic researchers should depend on their own valuation of biases and consider them when inferring any knowledge based on protein interaction networks. Besides the freely available database which predict the protein-protein interaction, commercially available software platforms such as Ingenuity Pathway Analysis (IPA) (<https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/>) and Metacore (<https://clarivate.com/products/metacore/>) are also considered great inclusive applications that enable analysis of many “omics” (Bessarabova et al. 2012; Yin et al. 2015) and agricultural proteomics as well; however, those software applications are mainly applied in medical proteomics rather than agricultural proteomics (Chen et al. 2013).

Proteomic analysis, in general, depends on data imaging which plays a serious role in understanding new results of proteomic research. In agricultural proteomic research especially for high-throughput experiments, heat maps are particularly suitable to achieve this mission, as they allow us to find measurable forms of result presentation across proteins concurrently. It is very useful to use heat maps for presenting comparative proteomic results organized in a simple yet expressive way. The superiority of a presented heat map can be highly improved by understanding and utilizing the options available in the online tools/software to organize the data in the heat map (Key 2012; Acton 2013). The idea of a heat map style of presentation appears to be originated from the use of color-based heat maps, which used to differentiate changes in temperatures. List of used websites/software to create heat maps used in proteomic research is listed in Table 1.5.

Table 1.5 Heat map generating tools/software/website

Heatmapper	Freely available web server that allows users to interactively visualize their data in the form of heat maps	http://www.heatmapper.ca	(Babicki et al. 2016)
ComplexHeat map	Software allow users to customize heatmaps	https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html	Gu et al. (2016)
InCHILB	Open source interactive JavaScript	https://openscreen.cz/software/inchlib/home/	Škuta et al. (2014)
InfernoRDN	Multi-omics heat map generation	https://omics.pnl.gov/software/infernordn	Sadler and Wright (2015)
Clustergrammer	Web-based tool for visualizing high-dimensional data	https://www.npmjs.com/package/clustergrammer	Fernandez et al. (2017)

1.5 Proteoinformatics of Edible Mushroom

Information regarding the life cycles and metabolisms of edible mushroom is of high importance for designing workable, fruitful, and effective cultivation process, especially with fungal species that are hard to propagate and need a special medium, temperature, etc. (Zhang et al. 2014a). Research on edible mushrooms' physiological changes, growth stages, development, interactions with the environment, and contribution in human diet used several different approaches from cell biology, physiology, and chemistry to the current and multi-omic techniques such as genomics (Chen et al. 2016), transcriptomics (Fu et al. 2017), proteomics (Rahmad et al. 2014), and metabolomics (Pandohee et al. 2015). Recently, the availability of bioinformatics related many edible mushrooms species helped to conduct many proteomic researches, thanks to the availability of their genome sequencing (Shim et al. 2016; Yang et al. 2017) due to the high request for edible mushrooms and their importance in food industry, medicine, and healthcare (Yap et al. 2014).

The availability of genome sequencing for those edible mushrooms allow researchers to run genome-based proteomics (Yap et al. 2015), which provided esteemed information for initiating molecular-based markers that can be used to improve the quality and usage of edible fungi. Recently, the importance of applying proteomic platforms in edible mushroom research has been highlighted, especially with nutraceutical and medicinal application possibilities (Al-Obaidi 2016b). Mushroom genome sequences make it possible for researchers to conduct research on mushroom growth (Tang et al. 2016; Wang et al. 2013a), developmental stages (Rahmad et al. 2014; Yin et al. 2012), and higher fungi medicinal properties (Yap et al. 2014).

1.6 Proteoinformatics of Animal Breeding Programs

The final products of terrestrial (cattle, poultry, and sheep) rigorous animal agro-farming systems have conventionally been mainly meat and milk products, fish, and other products from the aquaculture segment where both gained importance in terms of capacity and nutritional properties. Fundamental proteomics can be considered a promising tool for the discovery of protein diagnostic biomarkers for different and animal product quality markers.

Recently, the interest in studying livestock animals having proteomic and metabolomic platforms have increased rapidly (Suravajhala et al. 2016). Biomarker development in chicken was identified for different research goals, while in dairy cattle, numerous potential biomarkers were detected for meat and milk production (Goldansaz et al. 2017; Ortea et al. 2016). In domestic livestock and animal proteomics, the database search identification method in general is not an issue, since a comprehensive database of protein sequences is most probably available, databases

such as MetasSecKB (<http://bioinformatics.yzu.edu/secretomes/animal/index.php>) can be considered as a good reference. On the other side, in the cases that the animal genome has not been sequenced or not complete, other approaches such as de novo peptide sequencing is usually used (Blakeley et al. 2011). Commonly in the absence of enough proteoinformation, search against a protein sequence from closely related organisms. Small differences in peptide sequence from the sample and the genome/proteome database entries may guide to a big difference in protein identities. This issue obscures proteome analysis for non-sequenced species and between different subspecies, where the difference in the amino acid sequence of proteins is highlighted possible (Ignatchenko et al. 2017). These approaches are considered significant bioinformatic challenges because there are several aspects that affect or add inconsistency to determine protein identities. The availability of sufficient proteoinformatics data, the study of protein identification and metabolomic changes research considered the source for building models of whole systems. Such systems will permit investigators to understand the function of the protein complex in response to disease and environmental changes (Romero-Rodríguez et al. 2014). In the animal breeding proteomic research, proteomics may help in the search of animal biomarkers and offer more accurate health measures for livestock, which are essential for improving the breeding program, disease resistance, stress tolerance, and environmental changes (Marco-Ramell et al. 2016).

1.7 Conclusion

This chapter has concentrated mainly on the application of software programs and databases of proteomics in agricultural sciences, where the organism with no or incomplete genomic sequence data makes the identification of proteins more challenging in comparison to those highly studied organisms. The power of multi-omic methods for high-throughput identification and characterization of candidate genes tends to be lost in non-model organisms due to the lack of sufficient biological information. It is likely that the availability and accessibility of more sequence in plant/fungi and other agricultural-related organisms will ease some of these difficulties by making genomic data available for many non-model organisms. However, proteomic studies accumulatively produce huge amounts of data. It is usually done collecting protein annotations from databases. Answering biological questions using these data is still a great challenge. In conclusion, key objectives for agricultural proteoinformatics include the encouragement of sequence submission and make it available to the public research community. Finally, proteoinformatic databases, software programs, and methods need to be designated and utilized in a better way. Many tools and databases are adapted from human and specifically medical-related examination systems, and these may not be perfect for the analysis of plant, fungal, and other related agricultural proteomic data.

References

- Acton QA (2013) Issues in bioengineering and bioinformatics: 2013 Edition. ScholarlyEditions, Atlanta
- Afzal AJ, Natarajan A, Saini N, Iqbal MJ, Geisler M, El Shemy HA, Mungur R, Willmitzer L, Lightfoot DA (2009) The nematode resistance allele at the *rhg1* locus alters the proteome and primary metabolism of soybean roots. *Plant Physiol* 151(3):1264
- Agarwal GK, Bourguignon J, Rolland N, Ephritikhine G, Ferro M, Jaquinod M, Alexiou KG, Chardot T, Chakraborty N, Jolivet P (2011) Plant organelle proteomics: collaborating for optimal cell function. *Mass Spectrom Rev* 30(5):772–853
- Ahmad P, Latef AAA, Rasool S, Akram NA, Ashraf M, Guzel S (2016) Role of proteomics in crop stress tolerance. *Front Plant Sci* 7:1336
- Alexander MM, Cilia M (2016) A molecular tug-of-war: global plant proteome changes during viral infection. *Curr Plant Biol* 5(Supplement C):13–24. <https://doi.org/10.1016/j.cpb.2015.10.003>
- Allahverdiyeva Y, Battchikova N, Brosché M, Fujii H, Kangasjärvi S, Mulo P, Mähönen AP, Nieminen K, Overmyer K, Salojärvi J (2015) Integration of photosynthesis, development and stress as an opportunity for plant biology. *New Phytol* 208(3):647–655
- Al-Obaidi RJ, Mohd-Yusuf Y, Razali N, Jayapalan JJ, Tey C-C, Md-Noh N, Junit MS, Othman YR, Hashim HO (2014) Identification of proteins of altered abundance in oil palm infected with *Ganoderma boninense*. *Int J Mol Sci* 15(3):5175. <https://doi.org/10.3390/ijms15035175>
- Al-Obaidi J, Saidi N, Usuldin S, Rahmad N, Zean NB, Idris A (2016a) Differential proteomic study of oil palm leaves in response to in vitro inoculation with pathogenic and non-pathogenic *Ganoderma* spp. *J Plant Pathol* 98(2):235–244
- Al-Obaidi JR (2016b) Proteomics of edible mushrooms: a mini-review. *Electrophoresis* 37(10):1257–1263. <https://doi.org/10.1002/elps.201600031>
- Armengaud J, Trapp J, Pible O, Geffard O, Chaumont A, Hartmann EM (2014) Non-model organisms, a species endangered by proteogenomics. *J Proteomics* 105:5–18
- Ashok NR, Aparna HS (2017) Empirical and bioinformatic characterization of buffalo (*Bubalus bubalis*) colostrum whey peptides & their angiotensin I-converting enzyme inhibition. *Food Chem* 228(Supplement C):582–594. <https://doi.org/10.1016/j.foodchem.2017.02.032>
- Assumpcao TI, Fontes W, Sousa MV, Ricart CAO (2005) Proteome analysis of nelore bull (*Bos taurus indicus*) seminal plasma. *Protein Pept Lett* 12(8):813–817. <https://doi.org/10.2174/0929866054864292>
- Attwood TK, Coletta A, Muirhead G, Pavlopoulou A, Philippou PB, Popov I, Romá-Mateo C, Theodosiou A, Mitchell AL (2012) The PRINTS database: a fine-grained protein sequence annotation and analysis resource—its status in 2012. *Database* 2012:bas019–bas019. <https://doi.org/10.1093/database/bas019>
- Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS (2016) Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* 44(Web Server issue):W147–W153. <https://doi.org/10.1093/nar/gkw419>
- Babujee L, Wurtz V, Ma C, Lueder F, Soni P, Van Dorsselaer A, Reumann S (2010) The proteome map of spinach leaf peroxisomes indicates partial compartmentalization of phyloquinone (vitamin K1) biosynthesis in plant peroxisomes. *J Exp Bot* 61(5):1441–1453
- Balmant KM, Parker J, Yoo M-J, Zhu N, Dufresne C, Chen S (2015) Redox proteomics of tomato in response to *Pseudomonas syringae* infection. *Hortic Res* 2:15043. <https://doi.org/10.1038/hortres.2015.43>. <https://www.nature.com/articles/hortres201543#supplementary-information>
- Barah P, Bones AM (2014) Multidimensional approaches for studying plant defence against insects: from ecology to omics and synthetic biology. *J Exp Bot* 66(2):479–493
- Bartlett A, Lewis J, Williams ML (2016) Generations of interdisciplinarity in bioinformatics. *New Genet Soc* 35(2):186–209. <https://doi.org/10.1080/14636778.2016.1184965>
- Bengtsson T, Weighill D, Proux-Wéra E, Levander F, Resjö S, Burra DD, Moushib LI, Hedley PE, Liljeroth E, Jacobson D, Alexandersson E, Andreasson E (2014) Proteomics and transcrip-

- tomics of the BABA-induced resistance response in potato using a novel functional annotation approach. *BMC Genomics* 15(1):315. <https://doi.org/10.1186/1471-2164-15-315>
- Benton D (1996) Bioinformatics — principles and potential of a new multidisciplinary tool. *Trends Biotechnol* 14(8):261–272. [https://doi.org/10.1016/0167-7799\(96\)10037-8](https://doi.org/10.1016/0167-7799(96)10037-8)
- Bern M, Kil YJ, Becker C (2002) Bionic: advanced peptide and protein identification software. In: *Current protocols in bioinformatics*. Wiley <https://doi.org/10.1002/0471250953.bi1320s40>
- Bessarabova M, Ishkin A, JeBailey L, Nikolskaya T, Nikolsky Y (2012) Knowledge-based analysis of proteomics data. *BMC Bioinforma* 13(16):S13. <https://doi.org/10.1186/1471-2105-13-s16-s13>
- Bhadauria V (2016) *Omics in plant disease resistance*. Caister Academic Press, Norwich
- Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42(W1):W252–W258. <https://doi.org/10.1093/nar/gku340>
- Bischoff R, Permentier H, Guryev V, Horvatovich P (2016) Genomic variability and protein species—improving sequence coverage for proteogenomics. *J Proteomics* 134:25–36
- Blakeley P, Wright JC, Hubbard SJ, Jones AR (2011) Bioinformatics in animal proteomics. In: *Methods in animal proteomics*, Wiley-Blackwell, pp 103–119. <https://doi.org/10.1002/9780470960660.ch5>
- Bolger ME, Weisshaar B, Scholz U, Stein N, Usadel B, Mayer KFX (2014) Plant genome sequencing—applications for crop improvement. *Curr Opin Biotechnol* 26(Supplement C):31–37. <https://doi.org/10.1016/j.copbio.2013.08.019>
- Bruhn R, Jennings SF (2007) A multidisciplinary bioinformatics minor. *ACM SIGCSE Bull* 39(1):348–352
- Burgess S (2004) Proteomics in the chicken: tools for understanding immune responses to avian diseases. *Poult Sci* 83(4):552–573
- Cai-xia Z, Yi T, Li-yi Z, Ze-ran Z, Pei-hua C (2017) Comparative proteomic analysis of apple branches susceptible and resistant to ring rot disease. *Eur J Plant Pathol* 148(2):329–341. <https://doi.org/10.1007/s10658-016-1092-6>
- Calderone A, Cesareni G (2012) mentha: the interactome browser. 2012 18. <https://doi.org/10.14806/ej.18.A.455> p. 128
- Canovas FM, Dumas-Gaudot E, Recorbet G, Jorin J, Mock HP, Rossignol M (2004) Plant proteome analysis. *Proteomics* 4(2):285–298
- Cantú MD, Mariano AG, Palma MS, Carrilho E, Wulff NA (2008) Proteomic analysis reveals suppression of bark chitinases and proteinase inhibitors in citrus plants affected by the citrus sudden death disease. *Phytopathology* 98(10):1084–1092. <https://doi.org/10.1094/PHYTO-98-10-1084>
- Carpentier SC, Panis B, Vertommen A, Swennen R, Sergeant K, Renaut J, Laukens K, Witters E, Samyn B, Devreese B (2008) Proteome analysis of non-model plants: a challenging but powerful approach. *Mass Spectrom Rev* 27(4):354–377
- Chandramouli K, Qian P-Y (2009) Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum Genomics Proteomics* 2009:239204. <https://doi.org/10.4061/2009/239204>
- Chang J-W, Zhou Y-Q, Ul Qamar TM, Chen L-L, Ding Y-D (2016) Prediction of protein–protein interactions by Evidence Combining Methods. *Int J Mol Sci* 17(11). <https://doi.org/10.3390/ijms17111946>
- Chatr-aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, O’Donnell L, Oster S, Theesfeld C, Sellam A, Stark C, Breitkreutz B-J, Dolinski K, Tyers M (2017) The BioGRID interaction database: 2017 update. *Nucleic Acids Res* 45(Database issue):D369–D379. <https://doi.org/10.1093/nar/gkw1102>
- Chen C, McGarvey PB, Huang H, Wu CH (2010) Protein bioinformatics infrastructure for the integration and analysis of multiple high-throughput “omics” data. *Adv Bioinforma* 2010:423589. <https://doi.org/10.1155/2010/423589>
- Chen J, Chen L, Shen B (2013) Identification of network biomarkers for cancer diagnosis. In: Wang X (ed) *Bioinformatics of human proteomics*. Springer Netherlands, Dordrecht, pp 257–275. https://doi.org/10.1007/978-94-007-5811-7_11

- Chen L, Gong Y, Cai Y, Liu W, Zhou Y, Xiao Y, Xu Z, Liu Y, Lei X, Wang G, Guo M, Ma X, Bian Y (2016) Genome sequence of the edible cultivated mushroom *lentinula edodes* (Shiitake) reveals insights into lignocellulose degradation. *PLoS One* 11(8):e0160336. <https://doi.org/10.1371/journal.pone.0160336>
- Cho WCS (2007) Proteomics technologies and challenges. *Genomics Proteomics Bioinformatics* 5(2):77–85. [https://doi.org/10.1016/S1672-0229\(07\)60018-7](https://doi.org/10.1016/S1672-0229(07)60018-7)
- Chu X-L, Feng M-G, Ying S-H (2016) Qualitative ubiquitome unveils the potential significances of protein lysine ubiquitination in hyphal growth of *Aspergillus nidulans*. *Curr Genet* 62(1):191–201. <https://doi.org/10.1007/s00294-015-0517-7>
- Cipriano AKAL, Gondim DMF, Vasconcelos IM, Martins JAM, Moura AA, Moreno FB, Monteiro-Moreira ACO, Melo JGM, Cardoso JE, Paiva ALS, Oliveira JTA (2015) Proteomic analysis of responsive stem proteins of resistant and susceptible cashew plants after *Lasiodiplodia theobromae* infection. *J Proteomics* 113:90–109. <https://doi.org/10.1016/j.jprot.2014.09.022>
- Clark GW, Dar V-u-N, Bezginov A, Yang JM, Charlebois RL, Tillier ERM (2011) Using coevolution to predict protein–protein interactions. In: Cagney G, Emili A (eds) *Network biology: methods and applications*. Humana Press, Totowa, pp 237–256. https://doi.org/10.1007/978-1-61779-276-2_11
- Collins FS, Green ED, Guttmacher AE, Guyer MS (2003) A vision for the future of genomics research. *Nature* 422(6934):835–847
- Cooper B, Campbell KB, Feng J, Garrett WM, Frederick R (2011) Nuclear proteomic changes linked to soybean rust resistance. *Mol BioSystems* 7(3):773–783
- Cottrell J (2005) Database searching for protein identification and characterization
- Coumans JV, Poljak A, Raftery MJ, Backhouse D, Pereg-Gerk L (2009) Analysis of cotton (*Gossypium hirsutum*) root proteomes during a compatible interaction with the black root rot fungus *Thielaviopsis basicola*. *Proteomics* 9(2):335–349
- Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10(4):1794–1805. <https://doi.org/10.1021/pr101065j>
- Cristoni S, Mazzuca S (2011) Bioinformatics applied to proteomics. In: *Systems and computational biology-bioinformatics and computational modeling*. IntechOpen Limited, London
- Dahal D, Pich A, Braun HP, Wydra K (2010) Analysis of cell wall proteins regulated in stem of susceptible and resistant tomato species after inoculation with *Ralstonia solanacearum*: a proteomic approach. *Plant Mol Biol* 73(6):643–658. <https://doi.org/10.1007/s11103-010-9646-z>
- Dangl JL, Jones JDG (2001) Plant pathogens and integrated defence responses to infection. *Nature* 411(6839):826–833
- Dennis ES, Ellis J, Green A, Llewellyn D, Morell M, Tabe L, Peacock WJ (2008) Genetic contributions to agricultural sustainability. *Philos Trans R Soc B: Biol Sci* 363(1491):591–609. <https://doi.org/10.1098/rstb.2007.2172>
- Di Carli M, Villani ME, Bianco L, Lombardi R, Perrotta G, Benvenuto E, Donini M (2010) Proteomic analysis of the plant–virus interaction in Cucumber Mosaic Virus (CMV) resistant transgenic tomato. *J Proteome Res* 9(11):5684–5697. <https://doi.org/10.1021/pr100487x>
- Disruption ahead for NCBI databases (2016) Matrix science. <http://www.matrixscience.com/blog/disruption-ahead-for-ncbi-databases.html>. Accessed 20 Oct 2017
- Doblas VG, Smakowska-Luzan E, Fujita S, Alassimone J, Barberon M, Madalinski M, Belkhadir Y, Geldner N (2017) Root diffusion barrier control by a vasculature-derived peptide binding to the SGN3 receptor. *Science* 355(6322):280
- Eklblom R, Wolf JBW (2014) A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl* 7(9):1026–1042. <https://doi.org/10.1111/eva.12178>
- Elvira MI, Galdeano MM, Gilardi P, García-Luque I, Serra MT (2008) Proteomic analysis of pathogenesis-related proteins (PRs) induced by compatible and incompatible interactions of pepper mild mottle virus (PMMoV) in *Capsicum chinense* L3 plants. *J Exp Bot* 59(6):1253–1265. <https://doi.org/10.1093/jxb/ern032>
- Fang X, Jost R, Finnegan PM, Barbeti MJ (2013) Comparative proteome analysis of the strawberry-fusarium oxysporum f. sp. fragariae pathosystem reveals early activation of defense responses as a crucial determinant of host resistance. *J Proteome Res* 12(4):1772–1788

- Fears R (2007) Commission on genetic resources for food and agriculture
- Fen W, Baoxing S, Xing Z, Yaotian M, Dengyun L, Na Z, Pengfei J, Qing S, Jingfei H, Deli Z (2016) Prediction and analysis of the protein-protein interaction networks for chickens, cattle, dogs, horses and rabbits. *Curr Bioinform* 11(1):131–142. <https://doi.org/10.2174/1574893611666151203221255>
- Fernandez NF, Gundersen GW, Rahman A, Grimes ML, Rikova K, Hornbeck P, Ma'ayan A (2017) Clustergrammer, a web-based heatmap visualization and analysis tool for high-dimensional biological data. *Sci Data* 4:170151. <https://doi.org/10.1038/sdata.2017.151>. <https://www.nature.com/articles/sdata2017151#supplementary-information>
- Fernández-Acero FJ, Colby T, Harzen A, Cantoral JM, Schmidt J (2009) Proteomic analysis of the phytopathogenic fungus *Botrytis cinerea* during cellulose degradation. *Proteomics* 9(10):2892–2902
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44(D1):D279–D285. <https://doi.org/10.1093/nar/gkv1344>
- Fletcher J, Bender C, Budowle B, Cobb WT, Gold SE, Ishimaru CA, Luster D, Melcher U, Murch R, Scherm H, Seem RC, Sherwood JL, Sobral BW, Tolin SA (2006) Plant pathogen forensics: capabilities, needs, and recommendations. *Microbiol Mol Biol Rev* 70(2):450–471. <https://doi.org/10.1128/MMBR.00022-05>
- Fondi M, Liò P (2015) Multi -omics and metabolic modelling pipelines: challenges and tools for systems microbiology. *Microbiol Res* 171(Supplement C):52–64. <https://doi.org/10.1016/j.micres.2015.01.003>
- Fu Y, Dai Y, Yang C, Wei P, Song B, Yang Y, Sun L, Zhang Z-W, Li Y (2017) Comparative transcriptome analysis identified candidate genes related to Bailinggu mushroom formation and genetic markers for genetic analyses and breeding. *Sci Rep* 7:9266. <https://doi.org/10.1038/s41598-017-08049-z>
- Garg H, Li H, Sivasithamparam K, Barbetti MJ (2013) Differentially expressed proteins and associated histological and disease progression changes in cotyledon tissue of a resistant and susceptible genotype of *Brassica napus* infected with *Sclerotinia sclerotiorum*. *PLoS One* 8(6):e65205. <https://doi.org/10.1371/journal.pone.0065205>
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31(13):3784–3788
- Goldansaz SA, Guo AC, Sajed T, Steele MA, Plastow GS, Wishart DS (2017) Livestock metabolomics and the livestock metabolome: a systematic review. *PLoS One* 12(5):e0177675
- Gómez-Vidal S, Salinas J, Tena M, Lopez-Llorca LV (2009) Proteomic analysis of date palm (*Phoenix dactylifera* L.) responses to endophytic colonization by entomopathogenic fungi. *Electrophoresis* 30(17):2996–3005. <https://doi.org/10.1002/elps.200900192>
- Gong F, Hu X, Wang W (2015) Proteomic analysis of crop plants under abiotic stress conditions: where to focus our research? *Front Plant Sci* 6:418
- Gore S, Sanz García E, Hendrickx PMS, Gutmanas A, Westbrook JD, Yang H, Feng Z, Baskaran K, Berrisford JM, Hudson BP, Ikegawa Y, Kobayashi N, Lawson CL, Mading S, Mak L, Mukhopadhyay A, Oldfield TJ, Patwardhan A, Peisach E, Sahní G, Sekharan MR, Sen S, Shao C, Smart OS, Ulrich EL, Yamashita R, Quesada M, Young JY, Nakamura H, Markley JL, Berman HM, Burley SK, Velankar S, Kleywegt GJ (2017) Validation of structures in the protein data bank. *Structure* 25(12):1916–1927. <https://doi.org/10.1016/j.str.2017.10.009>
- Gu H, Zhu P, Jiao Y, Meng Y, Chen M (2011) PRIN: a predicted rice interactome network. *BMC Bioinforma* 12(1):161
- Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32(18):2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>
- Guijun D, Weidong P, Gongshe L (2006) The analysis of proteome changes in sunflower seeds induced by N+ implantation. *J Biosci (Bangalore)* 31(2):247–253. <https://doi.org/10.1007/BF02703917>

- Gupta N, Tanner S, Jaitly N, Adkins JN, Lipton M, Edwards R, Romine M, Osterman A, Bafna V, Smith RD, Pevzner PA (2007) Whole proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Res* 17(9):1362–1377. <https://doi.org/10.1101/gr.6427907>
- Hamady M, Tom Hiu Tung C, Resing K, Cios KJ, Knight R (2005) Key challenges in proteomics and proteoinformatics. *IEEE Eng Med Biol Mag* 24(3):34–40. <https://doi.org/10.1109/MEMB.2005.1436456>
- Hily JM, García A, Moreno A, Plaza M, Wilkinson MD, Fereres A, Fraile A, García-Arenal F (2014) The relationship between host lifespan and pathogen reservoir potential: an analysis in the system *Arabidopsis thaliana*-cucumber mosaic virus. *PLoS Pathog* 10(11):e1004492. <https://doi.org/10.1371/journal.ppat.1004492>
- Hoogland C, Mostaguir K, Sanchez J-C, Hochstrasser DF, Appel RD (2007) 2D PAGE databases for proteins in human body fluids. In: Thongboonkerd V (ed) *Proteomics of human body fluids: principles, methods, and applications*. Humana Press, Totowa, pp 137–146. https://doi.org/10.1007/978-1-59745-432-2_7
- Hosur R, Peng J, Vinayagam A, Stelz U, Xu J, Perrimon N, Bienkowska J, Berger B (2012) A computational framework for boosting confidence in high-throughput protein-protein interaction datasets. *Genome Biol* 13(8):R76. <https://doi.org/10.1186/gb-2012-13-8-r76>
- Hu J, Rampitsch C, Bykova NV (2015) Advances in plant proteomics toward improvement of crop productivity and stress resistance. *Front Plant Sci* 6:209. <https://doi.org/10.3389/fpls.2015.00209>
- Hulo N, Bairoch A, Bulliard V, Cerutti L, Cuče BA, de Castro E, Lachaize C, Langendijk-Genevaux PS, Sigrist CJA (2008) The 20 years of PROSITE. *Nucleic Acids Res* 36(Database issue):D245–D249. <https://doi.org/10.1093/nar/gkm977>
- Hutchins JRA (2014) What's that gene (or protein)? Online resources for exploring functions of genes, transcripts, and proteins. *Mol Biol Cell* 25(8):1187–1201. <https://doi.org/10.1091/mbc.E13-10-0602>
- Ignatchenko A, Sinha A, Alfaro JA, Boutros PC, Kislinger T, Ignatchenko V (2017) Detecting protein variants by mass spectrometry: a comprehensive study in cancer cell-lines. *Genome Med* 9(1):62
- Imam J, Nitin M, Toppo NN, Mandal NP, Kumar Y, Variar M, Bandopadhyay R, Shukla P (2014) A comprehensive overview on application of bioinformatics and computational statistics in rice genomics toward an Amalgamated approach for improving acquaintance base. In: Kishor PBK, Bandopadhyay R, Suravajhala P (eds) *Agricultural bioinformatics*. Springer India, New Delhi, pp 89–107. https://doi.org/10.1007/978-81-322-1880-7_5
- Jayaswal PK, Dogra V, Shanker A, Sharma TR, Singh NK (2017) A tree of life based on ninety-eight expressed genes conserved across diverse eukaryotic species. *PLoS One* 12(9):e0184276. <https://doi.org/10.1371/journal.pone.0184276>
- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40(D1):D109–D114. <https://doi.org/10.1093/nar/gkr988>
- Kang Y, Burton L, Lau A, Tate S (2017) SWATH-ID: an instrument method which combines identification and quantification in a single analysis. *Proteomics* 17(10):1500522–n/a. <https://doi.org/10.1002/pmic.201500522>
- Katam K, Jones KA, Sakata K (2015a) Advances in proteomics and bioinformatics in agriculture research and crop improvement. *J Proteomics Bioinforma* 8(3):39
- Katam R, Chibanguza K, Latinwo LM, Smith D (2015b) Proteome biomarkers in xylem reveal Pierce's disease tolerance in grape. *J Proteomics Bioinforma* 8(9):217–224
- Kaul S, Koo HL, Jenkins J, Rizzo M, Rooney T, Tallon LJ, Feldblyum T, Nierman W, Benito MI, Lin X (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408(6814):796–815
- Kaur P, Jost R, Sivasithamparam K, Barbetti MJ (2011) Proteome analysis of the Albugo candida–*Brassica juncea* pathosystem reveals that the timing of the expression of defence-related genes

- is a crucial determinant of pathogenesis. *J Exp Bot* 62(3):1285–1298. <https://doi.org/10.1093/jxb/erq365>
- Ke T, Yu J, Dong C, Mao H, Hua W, Liu S (2015) ocsESTdb: a database of oil crop seed EST sequences for comparative analysis and investigation of a global metabolic network and oil accumulation metabolism. *BMC Plant Biol* 15(1):19. <https://doi.org/10.1186/s12870-014-0399-8>
- Keskin O, Tuncbag N, Gursoy A (2016) Predicting protein–protein interactions from the molecular to the proteome level. *Chem Rev* 116(8):4884–4909
- Key M (2012) A tutorial in displaying mass spectrometry-based proteomic data using heat maps. *BMC Bioinform* 13(Suppl 16):S10–S10. <https://doi.org/10.1186/1471-2105-13-S16-S10>
- Khan FA (2015) *Biotechnology fundamentals*, 2nd edn. CRC Press, Boca Raton, Florida, United States
- Khan AM, Tan TW, Schönbach C, Ranganathan S (2013) APBioNet—Transforming Bioinformatics in the Asia-Pacific Region. *PLoS Comp Biol* 9(10):e1003317. <https://doi.org/10.1371/journal.pcbi.1003317>
- Khan I, Chen Y, Dong T, Hong X, Takeuchi R, Mori H, Kihara D (2014) Genome-scale identification and characterization of moonlighting proteins. *Biol Direct* 9(1):30. <https://doi.org/10.1186/s13062-014-0030-9>
- Khazanov NA, Carlson HA (2013) Exploring the composition of protein-ligand binding sites on a large scale. *PLoS Comp Biol* 9(11):e1003321. <https://doi.org/10.1371/journal.pcbi.1003321>
- Khoa Pham T, Wright PC (2007) Proteomic analysis of *Saccharomyces cerevisiae*. *Expert Rev Proteomics* 4(6):793–813
- Koller A, Washburn MP, Lange BM, Andon NL, Deciu C, Haynes PA, Hays L, Schieltz D, Ulaszek R, Wei J, Wolters D, Yates JR (2002) Proteomic survey of metabolic pathways in rice. *Proc Natl Acad Sci* 99(18):11969–11974
- Koltai H, Volpin H (2003) Agricultural genomics: an approach to plant protection. *Eur J Plant Pathol* 109(2):101–108. <https://doi.org/10.1023/A:1022512914003>
- Komatsu S, Mock H-P, Yang P, Svensson B (2013) Application of proteomics for improving crop protection/artificial regulation. *Front Plant Sci* 4:522. <https://doi.org/10.3389/fpls.2013.00522>
- Kumar A, Pathak RK, Gupta SM, Gaur VS, Pandey D (2015) Systems biology for smart crops and agricultural innovation: filling the gaps between genotype and phenotype for complex traits linked with robust agricultural productivity and sustainability. *OMICS: J Integr Biol* 19(10):581–601
- Land M, Hauser L, Jun S-R, Nookaew I, Leuze MR, Ahn T-H, Karpinetz T, Lund O, Kora G, Wassenaar T, Poudel S, Ussery DW (2015) Insights from 20 years of bacterial genome sequencing. *Funct Integr Genomics* 15(2):141–161. <https://doi.org/10.1007/s10142-015-0433-4>
- Lande NV, Subba P, Barua P, Gayen D, Keshava Prasad TS, Chakraborty S, Chakraborty N (2017) Dissecting the chloroplast proteome of chickpea (*Cicer arietinum* L.) provides new insights into classical and non-classical functions. *J Proteomics* 165(Supplement C):11–20. <https://doi.org/10.1016/j.jprot.2017.06.005>
- Larson RL, Wintermantel WM, Hill A, Fortis L, Nunez A (2008) Proteome changes in sugar beet in response to Beet necrotic yellow vein virus. *Physiol Mol Plant Pathol* 72(1):62–72. <https://doi.org/10.1016/j.pmp.2008.04.003>
- Lawrence S, Parker J, Chen S (2016) Plant response to bacterial pathogens: a proteomics view. In: Salekdeh GH (ed) *Agricultural proteomics volume 2: environmental stresses*. Springer International Publishing, Cham, pp 203–225. https://doi.org/10.1007/978-3-319-43278-6_9
- Lee J, Feng J, Campbell KB, Scheffler BE, Garrett WM, Thibivilliers S, Stacey G, Naiman DQ, Tucker ML, Pastor-Corrales MA, Cooper B (2009) Quantitative proteomic analysis of bean plants infected by a virulent and avirulent obligate rust fungus. *Mol Cell Proteomics* 8(1):19–31. <https://doi.org/10.1074/mcp.M800156-MCP200>
- Lei D, Lin R, Yin C, Li P, Zheng A (2014) Global protein–protein interaction network of rice sheath blight pathogen. *J Proteome Res* 13(7):3277–3293. <https://doi.org/10.1021/pr500069r>
- Lery LMS, Hemery AS, Nogueira EM, von Krüger WMA, Bisch PM (2010) Quantitative proteomic analysis of the interaction between the endophytic plant-growth-promoting bacterium

- gluconacetobacter diazotrophicus and sugarcane. *Mol Plant Microbe Interact* 24(5):562–576. <https://doi.org/10.1094/MPMI-08-10-0178>
- Leung T, Poulin R (2008) Parasitism, commensalism, and mutualism: exploring the many shades of symbioses. *Vie Milieu* 58(2):107–115
- Li T, Gong L, Wang Y, Chen F, Gupta VK, Jian Q, Duan X, Jiang Y (2017) Proteomics analysis of *Fusarium proliferatum* under various initial pH during fumonisin production. *J Proteomics* 164(Supplement C):59–72. <https://doi.org/10.1016/j.jprot.2017.05.008>
- Liu W, Gray S, Huo Y, Li L, Wei T, Wang X (2015) Proteomic analysis of interaction between a plant virus and its vector insect reveals new functions of hemipteran cuticular protein. *Mol Cell Proteomics*: MCP 14(8):2229–2242. <https://doi.org/10.1074/mcp.M114.046763>
- Lodha TD, Hembram P, Nitile Tep JB (2013) Proteomics: a successful approach to understand the molecular mechanism of plant-pathogen interaction. *Am J Plant Sci* 04(06):15. <https://doi.org/10.4236/ajps.2013.46149>
- Ma B (2015) Novor: real-time peptide de novo sequencing software. *J Am Soc Mass Spectrom* 26(11):1885–1894. <https://doi.org/10.1007/s13361-015-1204-0>
- Marco-Ramell A, de Almeida AM, Cristobal S, Rodrigues P, Roncada P, Bassols A (2016) Proteomics and the search for welfare and stress biomarkers in animal production in the one-health context. *Mol Biosyst* 12(7):2024–2035
- Martens L (2011) Proteomics databases and repositories. In: Wu CH, Chen C (eds) *Bioinformatics for comparative proteomics*. Humana Press, Totowa, pp 213–227. https://doi.org/10.1007/978-1-60761-977-2_14
- Maurer MH (2016) Two-dimensional gel electrophoresis image analysis via dedicated software packages. In: Marengo E, Robotti E (eds) *2-D PAGE map analysis: methods and protocols*. Springer New York, New York, pp 55–65. https://doi.org/10.1007/978-1-4939-3255-9_3
- McGarvey PB, Huang H, Mazumder R, Zhang J, Chen Y, Zhang C, Cammer S, Will R, Odle M, Sobral B, Moore M, Wu CH (2009) Systems integration of biodefense omics data for analysis of pathogen-host interactions and identification of potential targets. *PLoS One* 4(9):e7162. <https://doi.org/10.1371/journal.pone.0007162>
- Mehta A, Brasileiro ACM, Souza DSL, Romano E, Campos MA, Grossi-de-Sá MF, Silva MS, Franco OL, Fragoso RR, Bevitore R, Rocha TL (2008) Plant–pathogen interactions: what is proteomics telling us? *FEBS J* 275(15):3731–3746. <https://doi.org/10.1111/j.1742-4658.2008.06528.x>
- Memišević V, Zavaljevski N, Pieper R, Rajagopala SV, Kwon K, Townsend K, Yu C, Yu X, DeShazer D, Reifman J, Wallqvist A (2013) Novel Burkholderia mallei virulence factors linked to specific host-pathogen protein interactions. *Mol Cell Proteomics*: MCP 12(11):3036–3051. <https://doi.org/10.1074/mcp.M113.029041>
- Merrill SA, Mazza A-M, Council NR (2006) Genomics, proteomics, and the changing research environment
- Mirzaei M, Wu Y, Handler D, Maher T, Pascovici D, Ravishankar P, Moghaddam MZ, Haynes PA, Salekdeh GH, Chick JM (2016) Applications of quantitative proteomics in plant research. In: *Agricultural proteomics volume 1*. Springer, Switzerland. pp 1–29
- Mochida K, Shinozaki K (2010) Genomics and bioinformatics resources for crop improvement. *Plant Cell Physiol* 51(4):497–523. <https://doi.org/10.1093/pcp/pcq027>
- Mochida K, Shinozaki K (2011) Advances in omics and bioinformatics tools for systems analyses of plant functions. *Plant Cell Physiol* 52(12):2017–2038. <https://doi.org/10.1093/pcp/pcr153>
- Moortheis S, Hall A, Wright CF (2013) Informatics and clinical genome sequencing: opening the black box. *Genet Med* 15(3):165–171
- Nesvizhskii AI (2014) Proteogenomics: concepts, applications, and computational strategies. *Nat Methods* 11(11):1114–1125. <https://doi.org/10.1038/nmeth.3144>
- Newell-McGloughlin M (2008) Nutritionally improved agricultural crops. *Plant Physiol* 147(3):939–953. <https://doi.org/10.1104/pp.108.121947>
- Nilsson T, Mann M, Aebersold R, Yates JR, Bairoch A, Bergeron JMM (2010) Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Methods* 7(9):681–685
- Novák J, Lemr K, Schug KA, Havlíček V (2015) CycloBranch: de novo sequencing of non-ribosomal peptides from accurate product ion mass spectra. *J Am Soc Mass Spectrom* 26(10):1780–1786. <https://doi.org/10.1007/s13361-015-1211-1>

- Ong Q, Nguyen P, Phuong Thao N, Le L (2016) Bioinformatics approach in plant genomic research. *Curr Genomics* 17(4):368–378
- Orgogozo V, Morizot B, Martin A (2015) The differential view of genotype–phenotype relationships. *Front Genet* 6:179. <https://doi.org/10.3389/fgene.2015.00179>
- Ortea I, O'Connor G, Maquet A (2016) Review on proteomics for food authentication. *J Proteomics* 147:212–225
- Padula PM, Berry JI, O'Rourke BM, Raymond BB, Santos J, Djordjevic SP (2017) A comprehensive guide for performing sample preparation and top-down protein analysis. *Proteomes* 5(2):11. <https://doi.org/10.3390/proteomes5020011>
- Pandohee J, Stevenson PG, Conlan XA, Zhou X-R, Jones OAH (2015) Off-line two-dimensional liquid chromatography for metabolomics: an example using *Agaricus bisporus* mushrooms exposed to UV irradiation. *Metabolomics* 11 (4):939–951. <https://doi.org/10.1007/s11306-014-0749-4>
- Park S, Gupta R, Krishna R, Kim ST, Lee DY, Hwang D-j, Bae S-C, Ahn I-P (2016) Proteome analysis of disease resistance against *Ralstonia solanacearum* in Potato Cultivar CT206-10. *Plant Pathol J* 32(1):25–32. <https://doi.org/10.5423/PPJ.OA.05.2015.0076>
- Parker J, Koh J, Yoo M-J, Zhu N, Feole M, Yi S, Chen S (2013a) Quantitative proteomics of tomato defense against *Pseudomonas syringae* infection. *Proteomics* 13(12–13):1934–1946. <https://doi.org/10.1002/pmic.201200402>
- Parker J, Koh J, Yoo MJ, Zhu N, Feole M, Yi S, Chen S (2013b) Quantitative proteomics of tomato defense against *Pseudomonas syringae* infection. *Proteomics* 13(12–13):1934–1946
- Pearson WR (2013) An introduction to sequence similarity (“Homology”) searching. *Curr Protoc Bioinform/*editorial board, Andreas D Baxevanis [et al] 0 3:10.1002/0471250953.bi0471250301s0471250942. <https://doi.org/10.1002/0471250953.bi0301s42>
- Pechanova O, Pechan T, Ozkan S, McCarthy FM, Williams WP, Luthe DS (2010) Proteome profile of the developing maize (*Zea mays* L.) rachis. *Proteomics* 10(16):3051–3055. <https://doi.org/10.1002/pmic.200900833>
- Pegg GF (1981) Chapter 7 – Biochemistry and physiology of pathogenesis. In: *Fungal Wilt diseases of plants*, Academic Press, Cambridge, Massachusetts, United States. pp 193–253. <https://doi.org/10.1016/B978-0-12-464450-2.50012-7>
- Pérez-Clemente RM, Vives V, Zandalinas SI, López-Climent MF, Muñoz V, Gómez-Cadenas A (2013) Biotechnological approaches to study plant responses to stress. *Biomed Res Int* 2013:654120. <https://doi.org/10.1155/2013/654120>
- Pieper U, Webb BM, Dong GQ, Schneidman-Duhovny D, Fan H, Kim SJ, Khuri N, Spill YG, Weinkam P, Hammel M, Tainer JA, Nilges M, Sali A (2014) ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res* 42(D1):D336–D346. <https://doi.org/10.1093/nar/gkt1144>
- Pirovani CP, Carvalho HAS, Machado RCR, Gomes DS, Alvim FC, Pomella AWV, Gramacho KP, Cascardo JCM, Pereira GAG, Micheli F (2008) Protein extraction for proteome analysis from cacao leaves and meristems, organs infected by *Moniliophthora perniciosa*, the causal agent of the witches' broom disease. *Electrophoresis* 29(11):2391–2401. <https://doi.org/10.1002/elps.200700743>
- Placzek S, Schomburg I, Chang A, Jeske L, Ulbrich M, Tillack J, Schomburg D (2017) BRENDA in 2017: new perspectives and new tools in BRENDA. *Nucleic Acids Res* 45(D1):D380–D388. <https://doi.org/10.1093/nar/gkw952>
- Porteus B, Kocharunchitt C, Nilsson RE, Ross T, Bowman JP (2011) Utility of gel-free, label-free shotgun proteomics approaches to investigate microorganisms. *Appl Microbiol Biotechnol* 90(2):407–416. <https://doi.org/10.1007/s00253-011-3172-z>
- Qian D, Tian L, Qu L (2015) Proteomic analysis of endoplasmic reticulum stress responses in rice seeds. *Sci Rep* 5:14255. <https://doi.org/10.1038/srep14255>. <https://www.nature.com/articles/srep14255#supplementary-information>
- Rabonatahiry N, Chao H, Guo L, Gan J, Xiang J, Yan M, Zhang L, Yu L, Li M (2017) Synteny analysis of genes and distribution of loci controlling oil content and fatty acid profile based on QTL alignment map in *Brassica napus*. *BMC Genomics* 18(1):776. <https://doi.org/10.1186/s12864-017-4176-6>

- Rahmad N, Al-Obaidi JR, Rashid NMN, Zean NB, Yusoff MHYM, Shaharuddin NS, Jamil NAM, Saleh NM (2014) Comparative proteomic analysis of different developmental stages of the edible mushroom *Termitomyces heimii*. *Biol Res* 47(1):30. <https://doi.org/10.1186/0717-6287-47-30>
- Rao VS, Srinivas K, Sujini GN, Kumar GNS (2014) Protein-protein interaction detection: methods and analysis. *Int J Proteomics* 2014:12. <https://doi.org/10.1155/2014/147648>
- Romero-Rodríguez MC, Pascual J, Valledor L, Jorrín-Novo J (2014) Improving the quality of protein identification in non-model species. Characterization of *Quercus ilex* seed and *Pinus radiata* needle proteomes by using SEQUEST and custom databases. *J Proteomics* 105:85–91
- Sadler NC, Wright AT (2015) Activity-based protein profiling of microbes. *Curr Opin Chem Biol* 24:139–144. <https://doi.org/10.1016/j.cbpa.2014.10.022>
- Salvato F, Havelund JF, Chen M, Rao RSP, Rogowska-Wrzesinska A, Jensen ON, Gang DR, Thelen JJ, Møller IM (2014) The potato tuber mitochondrial proteome. *Plant Physiol* 164(2):637
- Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D (2004) The database of interacting proteins: 2004 update. *Nucleic Acids Res* 32(Suppl_1):D449–D451. <https://doi.org/10.1093/nar/gkh086>
- Schmidt UG, Endler A, Schelbert S, Brunner A, Schnell M, Neuhaus HE, Marty-Mazars D, Marty F, Baginsky S, Martinoia E (2007) Novel tonoplast transporters identified using a proteomic approach with vacuoles isolated from cauliflower buds. *Plant Physiol* 145(1):216
- Schneider M, Tognolli M, Bairoch A (2004) The Swiss-Prot protein knowledgebase and ExPASy: providing the plant community with high quality proteomic data and tools. *Plant Physiol Biochem* 42(12):1013–1021. <https://doi.org/10.1016/j.plaphy.2004.10.009>
- Sheynkman GM, Shortreed MR, Cesnik AJ, Smith LM (2016) Proteogenomics: integrating next-generation sequencing and mass spectrometry to characterize human proteomic variation. *Annu Rev Anal Chem (Palo Alto, Calif)* 9(1):521–545. <https://doi.org/10.1146/annurev-anchem-071015-041722>
- Shim D, Park S-G, Kim K, Bae W, Lee GW, Ha B-S, Ro H-S, Kim M, Ryoo R, Rhee S-K, Nou I-S, Koo C-D, Hong CP, Ryu H (2016) Whole genome de novo sequencing and genome annotation of the world popular cultivated edible mushroom, *Lentinula edodes*. *J Biotechnol* 223(Supplement C):24–25. <https://doi.org/10.1016/j.jbiotec.2016.02.032>
- Singh R (2015) *Bioinformatics: genomics and proteomics*. Vikas Publishing House, Chennai, India
- Sinha R, Bhattacharyya D, Majumdar AB, Datta R, Hazra S, Chattopadhyay S (2013) Leaf proteome profiling of transgenic mint infected with *Alternaria alternata*. *J Proteomics* 93:117–132. <https://doi.org/10.1016/j.jprot.2013.01.020>
- Škuta C, Bartůněk P, Svozil D (2014) InChIlib – interactive cluster heatmap for web applications. *J Cheminformatics* 6(1):44. <https://doi.org/10.1186/s13321-014-0044-4>
- Song F, Qi D, Liu X, Kong X, Gao Y, Zhou Z, Wu Q (2015) Proteomic analysis of symbiotic proteins of *Glomus mosseae* and *Amorpha fruticosa*. *Sci Rep* 5:18031. <https://doi.org/10.1038/srep18031>. <https://www.nature.com/articles/srep18031#supplementary-information>
- Stare T, Stare K, Weckwerth W, Wienkoop S, Gruden K (2017) Comparison between proteome and transcriptome response in potato (*Solanum tuberosum* L.) leaves following Potato virus Y (PVY) infection. *Proteomes* 5(3):14
- Subramanian S, Cho U-H, Keyes C, Yu O (2009) Distinct changes in soybean xylem sap proteome in response to pathogenic and symbiotic microbe interactions. *BMC Plant Biol* 9(1):119. <https://doi.org/10.1186/1471-2229-9-119>
- Suravajhala P, Kogelman LJ, Kadarmideen HN (2016) Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare. *Genet Sel Evol* 48(1):38
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43(Database issue):D447–D452. <https://doi.org/10.1093/nar/gku1003>
- Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P (2017a) The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res* 45(D1):D362–D368

- Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C (2017b) The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res* 45(Database issue):D362–D368. <https://doi.org/10.1093/nar/gkw937>
- Taheri F, Nematzadeh G, Zamharir MG, Nekouei MK, Naghavi M, Mardi M, Salekdeh GH (2011) Proteomic analysis of the Mexican lime tree response to “Candidatus *Phytoplasma aurantifolia*” infection. *Mol BioSystems* 7(11):3028–3035. <https://doi.org/10.1039/c1mb05268c>
- Tamburino R, Vitale M, Ruggiero A, Sassi M, Sannino L, Arena S, Costa A, Batelli G, Zambrano N, Scaloni A, Grillo S, Scotti N (2017) Chloroplast proteome response to drought stress and recovery in tomato (*Solanum lycopersicum* L.). *BMC Plant Biol* 17(1):40. <https://doi.org/10.1186/s12870-017-0971-0>
- Tan BC, Lim YS, Lau S-E (2017) Proteomics in commercial crops: an overview. *J Proteomics*. <https://doi.org/10.1016/j.jprot.2017.05.018>
- Tang LH, Tan Q, Bao DP, Zhang XH, Jian HH, Li Y, Yang R, Wang Y (2016) Comparative proteomic analysis of light-induced mycelial Brown film formation in *Lentinula edodes*. *Biomed Res Int* 2016:8. <https://doi.org/10.1155/2016/5837293>
- The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45(Database issue):D158–D169. <https://doi.org/10.1093/nar/gkw1099>
- Thrall PH, Oakeshott JG, Fitt G, Southerton S, Burdon JJ, Sheppard A, Russell RJ, Zalucki M, Heino M, Ford Denison R (2011) Evolution in agriculture: the application of evolutionary approaches to the management of biotic interactions in agro-ecosystems. *Evol Appl* 4(2):200–215. <https://doi.org/10.1111/j.1752-4571.2010.00179.x>
- Tyanova S, Temu T, Cox J (2016a) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11:2301. <https://doi.org/10.1038/nprot.2016.136>
- Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016b) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13:731. <https://doi.org/10.1038/nmeth.3901>. <https://www.nature.com/articles/nmeth.3901#supplementary-information>
- Van De Wouw AP, Howlett BJ (2011) Fungal pathogenicity genes in the age of ‘omics’. *Mol Plant Pathol* 12(5):507–514
- Van Emon JM (2016) The omics revolution in agricultural research. *J Agric Food Chem* 64(1):36–44. <https://doi.org/10.1021/acs.jafc.5b04515>
- Vanderschuren H, Lentz E, Zainuddin I, Gruissem W (2013) Proteomics of model and crop plant species: status, current limitations and strategic advances for crop improvement. *J Proteomics* 93:5–19
- Vu LD, Stes E, Van Bel M, Nelissen H, Maddelein D, Inzé D, Coppens F, Martens L, Gevaert K, De Smet I (2016) Up-to-date workflow for plant (Phospho)proteomics identifies differential drought-responsive phosphorylation events in Maize leaves. *J Proteome Res* 15(12):4304–4317. <https://doi.org/10.1021/acs.jproteome.6b00348>
- Wang F-X, Ma Y-P, Yang C-L, Zhao P-M, Yao Y, Jian G-L, Luo Y-M, Xia G-X (2011) Proteomic analysis of the sea-island cotton roots infected by wilt pathogen *Verticillium dahliae*. *Proteomics* 11(22):4296–4309. <https://doi.org/10.1002/pmic.201100062>
- Wang M, Gu B, Huang J, Jiang S, Chen Y, Yin Y, Pan Y, Yu G, Li Y, Wong BHC, Liang Y, Sun H (2013a) Transcriptome and proteome exploration to provide a resource for the study of *Agrocybe aegerita*. *PLoS One* 8(2):e56686. <https://doi.org/10.1371/journal.pone.0056686>
- Wang Y, Kim SG, Wu J, Huh HH, Lee SJ, Rakwal R, Kumar Agrawal G, Park ZY, Young Kang K, Kim ST (2013b) Secretome analysis of the rice bacterium *Xanthomonas oryzae* (Xoo) using in vitro and in planta systems. *Proteomics* 13(12–13):1901–1912
- Webb K, Broccardo C, Prenni J, Wintermantel W (2014) Proteomic profiling of sugar beet (*Beta vulgaris*) leaves during rhizomania compatible interactions. *Proteomes* 2(2):208
- Wei Z, Wang Z, Li X, Zhao Z, Deng M, Dong Y, Cao X, Fan G (2017) Comparative proteomic analysis of *Paulownia fortunei* response to phytoplasma infection with dimethyl sulfate treatment. *Int J Genomics* 2017:11. <https://doi.org/10.1155/2017/6542075>

- Wenger CD, Phanstiel DH, Lee MV, Bailey DJ, Coon JJ (2011) COMPASS: a suite of pre- and post-search proteomics software tools for OMSSA. *Proteomics* 11(6):1064–1074. <https://doi.org/10.1002/pmic.201000616>
- Wu CH, Yeh L-SL, Huang H, Arminski L, Castro-Alvares J, Chen Y, Hu Z, Kourtesis P, Ledley RS, Suzek BE (2003) The protein information resource. *Nucleic Acids Res* 31(1):345–347
- Wu L, Wang S, Wu J, Han Z, Wang R, Wu L, Zhang H, Chen Y, Hu X (2015) Phosphoproteomic analysis of the resistant and susceptible genotypes of maize infected with sugarcane mosaic virus. *Amino Acids* 47(3):483–496. <https://doi.org/10.1007/s00726-014-1880-2>
- Xiong J (2006) *Essential bioinformatics*. Cambridge University Press, United Kingdom
- Xu H, Freitas MA (2009) MassMatrix: a database search program for rapid characterization of proteins and peptides from Tandem Mass Spectrometry Data. *Proteomics* 9(6):1548–1555. <https://doi.org/10.1002/pmic.200700322>
- Xue Y, Ren J, Gao X, Jin C, Wen L, Yao X (2008) GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. *Mol Cell Proteomics: MCP* 7(9):1598–1608. <https://doi.org/10.1074/mcp.M700574-MCP200>
- Yang R, Li Y, Song X, Tang L, Li C, Tan Q, Bao D (2017) The complete mitochondrial genome of the widely cultivated edible fungus *Lentinula edodes*. *Mitochondrial DNA Part B* 2(1):13–14. <https://doi.org/10.1080/23802359.2016.1275839>
- Yao YA, Wang J, Ma X, Lutts S, Sun C, Ma J, Yang Y, Achal V, Xu G (2012) Proteomic analysis of Mn-induced resistance to powdery mildew in grapevine. *J Exp Bot* 63(14):5155–5170. <https://doi.org/10.1093/jxb/ers175>
- Yap H-YY, Chooi Y-H, Firdaus-Raih M, Fung S-Y, Ng S-T, Tan C-S, Tan N-H (2014) The genome of the Tiger Milk mushroom, *Lignosus rhinocerotis*, provides insights into the genetic basis of its medicinal properties. *BMC Genomics* 15(1):635. <https://doi.org/10.1186/1471-2164-15-635>
- Yap H-YY, Fung S-Y, Ng S-T, Tan C-S, Tan N-H (2015) Genome-based proteomic analysis of *Lignosus rhinocerotis* (Cooke) Ryvarden sclerotium. *Int J Med Sci* 12(1):23
- Yates JI JR, Gilchrist A, Howell KE, Bergeron JJM (2005) Proteomics of organelles and large cellular structures. *Nat Rev Mol Cell Biol* 6:702. <https://doi.org/10.1038/nrm1711>
- Yin Y, Yu G, Chen Y, Jiang S, Wang M, Jin Y, Lan X, Liang Y, Sun H (2012) Genome-wide transcriptome and proteome analysis on different developmental stages of *Cordyceps militaris*. *PLoS One* 7(12):e51853
- Yin S-Y, Pradeep MS, Yang N-S (2015) Use of omics approaches for developing immune-modulatory and anti-inflammatory phyto-medicines. In: *Genomics, proteomics and metabolomics in nutraceuticals and functional foods*. Wiley, pp 453–475. <https://doi.org/10.1002/9781118930458.ch36>
- Yun Z, Gao H, Liu P, Liu S, Luo T, Jin S, Xu Q, Xu J, Cheng Y, Deng X (2013) Comparative proteomic and metabolomic profiling of citrus fruit with enhancement of disease resistance by postharvest heat treatment. *BMC Plant Biol* 13(1):44. <https://doi.org/10.1186/1471-2229-13-44>
- Zhan X, Long Y, Lu M (2017) Exploration of variations in proteome and metabolome for predictive diagnostics and personalized treatment algorithms: Innovative approach and examples for potential clinical application. *J Proteomics*. <https://doi.org/10.1016/j.jprot.2017.08.020>
- Zhang Y, Geng W, Shen Y, Wang Y, Dai Y-C (2014a) Edible mushroom cultivation for food security and rural development in China: bio-innovation, technological dissemination and marketing. *Sustainability* 6(5):2961–2973
- Zhang Y, Nandakumar R, Bartelt-Hunt SL, Snow DD, Hodges L, Li X (2014b) Quantitative proteomic analysis of the *Salmonella*-lettuce interaction. *Microbial Biotechnol* 7(6):630–637. <https://doi.org/10.1111/1751-7915.12114>
- Zhang C-x, Tian Y, Cong P-h (2015) Proteome analysis of pathogen-responsive proteins from apple leaves induced by the *Alternaria* Blotch *Alternaria alternata*. *PLoS One* 10(6):e0122233. <https://doi.org/10.1371/journal.pone.0122233>
- Zheng A, Luo J, Meng K, Li J, Zhang S, Li K, Liu G, Cai H, Bryden WL, Yao B (2014) Proteome changes underpin improved meat quality and yield of chickens (*Gallus gallus*) fed the probiotic *Enterococcus faecium*. *BMC Genomics* 15(1):1167. <https://doi.org/10.1186/1471-2164-15-1167>

- Zheng A, Luo J, Meng K, Li J, Bryden WL, Chang W, Zhang S, Wang LXN, Liu G, Yao B (2016) Probiotic (*Enterococcus faecium*) induced responses of the hepatic proteome improves metabolic efficiency of broiler chickens (*Gallus gallus*). *BMC Genomics* 17(1):89. <https://doi.org/10.1186/s12864-016-2371-5>
- Zhu P, Gu H, Jiao Y, Huang D, Chen M (2011) *Genomics Proteomics Bioinformatics* 9(4):128–137. [https://doi.org/10.1016/S1672-0229\(11\)60016-8](https://doi.org/10.1016/S1672-0229(11)60016-8)
- Zhu G, Wu A, Xu X-J, Xiao P, Lu L, Liu J, Cao Y, Chen L, Wu J, Zhao X-M (2015) PPIM: a protein-protein interaction database for maize. *Plant Physiol.* pp. 01821.02015
- Zhu Y, Engström PG, Tellgren-Roth C, Baudo CD, Kennell JC, Sun S, Billmyre RB, Schröder MS, Andersson A, Holm T, Sigurgeirsson B, Wu G, Sankaranarayanan SR, Siddharthan R, Sanyal K, Lundeberg J, Nystedt B, Boekhout T, Dawson TL, Heitman J, Scheynius A, Lehtö J (2017) Proteogenomics produces comprehensive and highly accurate protein-coding gene annotation in a complete genome assembly of *Malassezia sympodialis*. *Nucleic Acids Res* 45(5):2629–2643. <https://doi.org/10.1093/nar/gkx006>