Current publications of interest

**Microalgae**

*Current Research and Applications*
Edited by: MN Tsaloglou
152 pp, January 2016
The latest research and newest approaches to the study of microalgae.

**Bacteria-Plant Interactions**

*Advanced Research and Future Trends*
Edited by: J Murillo, BA Vinatzer, RW Jackson, et al.
x + 228 pp, March 2015
"an up-to-date overview" (Ringgold)

**Microarrays**

*Current Technology, Innovations and Applications*
Edited by: Z He
x + 246 pp, August 2014
"a valuable and useful source ... recommended" (Biotechnol. Agron. Soc. Environ.)

**Metagenomics of the Microbial Nitrogen Cycle**

*Theory, Methods and Applications*
Edited by: D Marco
xv + 268 pp, September 2014
"a strong overview" (Ringgold)

**Proteomics**

*Targeted Technology, Innovations and Applications*
Edited by: M Fuentes, J LaBaer
x + 186 pp, September 2014
"many excellent chapters" (Doody's)

**Applied RNAi**

*From Fundamental Research to Therapeutic Applications*
Edited by: P Arbuthnot, MS Weinberg
x + 252 pp, June 2014
"Essential reading" (Biotechnol Agron Soc Environ); "recommended" (Fungal Diversity); "an excellent resource" (Doody's)

**Next-generation Sequencing**

*Current Technologies and Applications*
Edited by: J Xu
xii + 160 pp, March 2014
Book: ISBN 978-1-908230-3-3 £120/$240
"written in an accessible style" (Zentralblatt Math); "recommend this book to all investigators" (ChemMedChem)

**Omics in Soil Science**

*Edited by: P Nannipieri, G Pietramellara, G Renella*
x + 198 pp, January 2014
"a recommended reference" (Biotechnol. Agrom. Soc. Environ.); "a must for Soil scientists" (Fungal Diversity)

**Genome Analysis**

*Current Procedures and Applications*
Edited by: MS Poptsova
xv + 374 pp, January 2014
An up-to-date and comprehensive overview of next-generation sequencing data analysis, highlighting problems and limitations, applications and developing trends in various fields of genome research.

**RNA Editing**

*Current Research and Future Trends*
Edited by: S Maas
viii + 240 pp, June 2013
"an essential book" (Doody's)

**Real-Time PCR**

*Advanced Technologies and Applications*
Edited by: NA Saunders, MA Lee
viii + 284 pp, July 2013
"an invaluable reference" (Doody's); "wide range of real time PCR technologies" (Food Sci Technol Abs); "I was impressed by this text" Aus J Med Sci

**Bionanotechnology**

*Biological Self-assembly and its Applications*
Edited by: BHA Rehm
x + 310 pp, February 2013
"the most striking and successful approaches" Book News
Linking Biomarker and Comparative Omics to Pathogens in Legumes

Marwan Diapari
Southern Crop Protection and Food Research Centre, Agriculture and Agri-Food Canada / Agriculture et Agroalimentaire Canada, 1391 Sandford St., London, ON, N5V 4T3 Canada. Email Marwan.Diapari@agr.gc.ca
DOI: http://dx.doi.org/10.21775/9781910190357.13

Abstract
It is envisioned that a more precise study of the association between the traits and biomarkers will dramatically decrease the time and costs required to bring new improved disease resistance lines to market. The field of omics has an enormous potential to assess diseases more precise, including the identification and understanding of pathogenic mechanisms in legume crops, and have been exemplified by a relatively large number of studies. Recently, molecular genetic studies have accumulated a huge amount of genotypic data, through a more affordable next generation sequencing (NGS) technology, causing the omics approaches to fall behind. In this paper I provide an overview of genomics and proteomics and their use in legume crops, including the use of comparative genomics to identify homologous markers within legume crops.

Introduction
Legumes are economically important, rich source of quality protein for humans and animals, and enrich the soil by producing their own nitrogen in symbiosis with nitrogen-fixing bacteria (Singh et al., 2007). A better knowledge of its physiological characteristics and behaviour within its own environment is imperative in order to increase legume competitiveness and to adapt its production to the new constraints arising from the climatic changes and the occurrence of new diseases (Adam-Blondon, 2007).

Molecular approaches (OMICS) that involve mRNA (transcriptomics), proteins (proteomics), metabolites (metabolomics), are indispensable in the era of functional genomics, and have been exemplified by a relatively large number of studies, such as gene expression, genetic diversity, phylogeny, comparative genomics and epigenetics (Ma et al., 2013; Sabir et al., 2014). OMICS has brought the possibilities for genetic improvement, conservation of biodiversity, and imparting disease resistance in crops for further targeted breeding (Rubiales et al., 2015). Maintaining genetic diversity and capturing genetic improvement are often through structuring populations into a genetic hierarchy, setting population sizes, and managing pedigrees and inbreeding.

Plant diseases can drastically abate the crop yields as the degree of disease outbreak is getting severe around the world. Major diseases are rusts, powdery and downy mildews, ascochyta blights, botrytis gray molds, anthracnose, damping-off, root rots, collar rots, vascular wilts and white molds (Gururani et al., 2012; Rubiales et al., 2015). New disease occurrence can involve gene duplication, transfer or recombination (Holomone and Brent, 2009). The negative impact of diseases and the genetic mechanisms and gene interactions involved in the resistance response are still poorly understood (Cantu et al., 2013). Comprehensive body of knowledge and understanding of the cellular functions will aid the development of novel practical approaches to enhance disease resistance (Choi and Hwang, 2014).

Legume crops have been exploited by breeding program (Robertson et al., 1996), including for disease purpose breeding program. It is believed that wild relatives are considered as important source of genetic diversity for agriculture, although their gene repertoire remains largely unexplored (Li et al., 2014). Proposed draft legume genomes allow the prediction of the full complement of genes, phylogenetic relationships and chromosomal locations will provide breeders and scientists with a useful tool to identify novel disease resistance traits (Andolfo et al., 2014). It is expected the array will accelerate hypothesis generation and gene discovery in disease defense pathways, responses to abiotic stresses, development, and evolutionary diversity in plants (Close et al., 2004). Directed molecular evolution can be used to develop enabling technologies for plant biologists; how genes can be optimized to generate improved input traits such as those conferring insect tolerance, disease control and herbicide tolerance; and how plant quality can be altered to improve yield (Lassner and McElroy, 2002). Molecular genetics provide an unparalleled opportunity to understand how genes and genetic changes interact with environmental stimuli to either preserve health or cause disease (Schwartz et al., 2004).

Plant breeders strive to stay ahead of the evolving pathogens by releasing cultivars with new resistance genes or gene combinations (Webb and Fellers, 2006). Breeding for disease resistance in crops has mainly been accomplished by incorporating single resistance gene. There are advantages to quantitative resistance in terms of durability. New technologies in genomics and proteomics are providing insights into disease-resistance pathways (Jordan et al., 2006). Genetic improvement of the crop is being actively pursued and numerous functional genomics studies aimed at characterizing gene controlling key agronomic characteristics and providing some insight into the basis of disease resistances (Hulbert and Pumphrey, 2014; Mellor et al., 2012).
The fields of environmental genetics and environmental genomics has enormous potential to affect our ability to accurately assess the risk of developing disease, identify and understand basic pathogenic mechanisms that are critical to disease progression (Schwartz et al., 2004). Genetic studies pathogen during the past decade have made an excellent system for investigating pathogen-host interactions (Xu and Xue, 2002).

Genetic variation has an important role in determining biologic phenotypes and disease resistance (Sladek and Hudson, 2006). Inter-genomic comparisons identified lineage-specific genes, some of which show evidence of positive selection and may contribute to variation of agronomic traits such as disease resistance (Li et al., 2014). Scarcity of genetic polymorphism makes crops especially vulnerable to a wide variety of pathogens (Ratnaparkhe et al., 2011). Sequence analysis of plant disease resistance genes shows similarity among themselves (Ramachandra et al., 2011) and are present in abundance and majority of these encode proteins with a nucleotide-binding site (NBS) (Basak et al., 2007; Ribas et al., 2011), comprising up to 2% of all genes (Mun et al., 2009).

NBS encoding resistance genes are key plant disease-resistance genes (R-genes) represent the major class of disease resistance genes (Mun et al., 2009; Richly et al., 2002) also play an important role in offering resistance to pathogens (Yu et al., 2014). Numerous plant R-genes have been used with varying degree of success in crop improvement (Gururani et al., 2012). A number of studies have focused on NBS-encoding genes in disease resistant breeding programs for diverse plants (Wei et al., 2013). Plant possesses a large number of R-genes with diverse recognition specificities which are activated in response to a variety of microbial pathogens (Sharma et al., 2014) and have the ability to detect a pathogen attack and facilitate a counter attack against the pathogen (Gururani et al., 2012; Martin et al., 2003) also often show significant similarity amongst themselves in terms of both their DNA sequences and structural motifs present in their protein products (López et al., 2003).

The majority of verified plant disease resistance genes isolated to date are of the NBS-LRR class (Madsen et al., 2003). The NBS-LRR genes are the largest class of disease resistance genes in plants (Wei et al., 2013; Xu et al., 2011). NBS profiling, in particular, targets conserved nucleotide binding site-encoding sequences of resistance gene analogs (RGAs), and is widely used to identify molecular markers for R genes (Vossen et al., 2013).

Reactive oxygen species (ROS) signaling network mediating disease resistance in plants information became available, thanks to the complete Arabidopsis (Arabidopsis thaliana) and rice (Oryza sativa) genome sequences (Kotchoni and Gachomo, 2006). ROS play multiple roles in interactions between plants and microbes (Daub et al., 2013) and in the plant hypersensitive disease resistance response (Cecconi et al., 2009), including induced systemic resistances (Deepak et al., 2008).

Genomic and Proteomics
Recent advances in metabolomics, genomics, transcriptomic and metabolic modelling offer new opportunities to address this question and generate a system-level understanding of metabolic interactions at the host-pathogen interface (Rico et al., 2011). Functional genomics represents a systematic approach to elucidating the function of the novel genes revealed by complete genome sequences. The different approaches are grouped into four domains: genome, transcriptome, proteome, and metabolome (Oliver, 2002). Genomics is a new and promising area, which can broadly be defined as the application of high throughputs genomics and functional genomic technologies to the study of pathology in plants (Brown and van der Ouderaa, 2007). Access to complete genomic sequences, coupled with rapidly accumulating data related to RNA and protein expression patterns, has made it possible to determine comprehensively how genes contribute to complex phenotypes (Close et al., 2004).

The huge amount of genomic data now becoming available offers both opportunities and challenges for breeding program (Garrett et al., 2006). Innovative approaches may be required to sequence, assemble, annotate and analyse host and pathogen genomes. Overcoming these challenges will enable scientists to investigate the genes and genome organisation may ultimately lead to new solutions for control diseases legume (Van Zee et al., 2007). The methods utilised provide new opportunities for studying the nature and role of defence mechanisms in plants (Collinge et al., 2008).

Knowledge of structural and functional genomics and genes and comparative genomics with model species has increased our understanding genomic diversity through integration of genetical, evolutionary and structural data (Heslop-Harrison and Schwarzacher, 2007). We will also gain a better understanding of the genes responsible for disease resistance. With a collection of desirable genes in mind, we can again use genomics as a diagnostic tool to search for these genes in the wide variety and follow their transfer by classical breeding (Regan et al., 2006). Following largescale source of sequencing of libraries and bioinformatics searches in public databases pave the way for the design and conduct of large-scale experiments that will help to a better understanding the complex biology of disease resistance (Bernier, 2004) and facilitating the isolation of novel genes, as well as by helping to identify targets for chemical control (Kamoun et al., 2002).

DNA sequencing is now complete for some hosts and several pathogens to form predictions about epidemic features and outcomes and for understanding host resistance and pathogen evolution (Garrett et al., 2006). The use of large-scale complementary DNA library constructions and genome-wide transcript profiles of plants exposed to biotic stress provide the data required to drive hypotheses concerning the function of newly identified genes (Collinge et al., 2008). For example, an extensive programme of functional genomics is being undertaken through the systematic analysis of insertionional mutants in Arabidopsis (Delseny and Pelletier, 2001).
Proteins are involved in disease resistance and defense (Mooney et al., 2006). In recent years, proteomics has been used to gain in-depth understanding of many aspects of the host defense against pathogens to identify differentially accumulated proteins a mutant and a wild type (Chen et al., 2013; Zimaro et al., 2011). Fueled by ever growing DNA sequence information, proteomics is providing insight into the mechanisms of biological processes in a high-throughput mode (Sharma et al., 2008). Proteomic approach was employed to investigate the cold stress-responsive proteins in citrus (Long et al., 2012). Various proteomic techniques for isolation and purification of proteins have been developed; each tailored to preserve protein properties relevant to study of a particular disease type (Vo and Palsson, 2007). The application of combined proteomic and RNA interfering analyses is an efficient strategy to identify genes required disease resistance and probably other biological processes in plants (Xu et al., 2012).

The recent development of large scale phenotyping, genome sequencing and analysis of gene, protein and metabolite expressions can be of great help to further decipher plant-pathogen interactions and identify key resistance components that may be introgressed into crop plants through breeding (Rubiales et al., 2015). With the sequencing of entire genomes it has become technically feasible to study transcription on a global scale (Barnett and Fisher, 2006). In 2001, The Medicago Genome Initiative (MGI) is a database of EST sequences of the model legume Medicago truncatula (http://www.noble.org), which is taking a global approach in studying the genetic and biochemical events associated with the growth, development and environmental interactions of this model legume (Bell et al., 2001). Genome sequencing of the model legumes, Medicago truncatula and Lotus japonicus provides an opportunity for large-scale sequence-based comparison of two genomes in the same plant family (Cannon et al., 2006). Recent genome sequencing enables mega-base scale comparisons between related genomes. In 2013, chickpea draft sequence genome using next-generation sequencing platforms, bacterial artificial chromosome end sequences and a genetic map has been proposed. The draft genome sequence is expected to facilitate genetic enhancement and breeding to develop improved chickpea varieties (Jain et al., 2013). In the same year whole genome shotgun sequence of CDC Frontier, a kabuli chickpea variety, this contains an estimated 28,269 genes published. Candidate genes for disease resistance and agronomic traits are highlighted, including traits that distinguish the two main market classes of cultivated chickpea--desi and kabuli. These data comprise a resource for chickpea improvement through molecular breeding and provide insights into both genome diversity and domestication (Varshney et al., 2013). The genome of V. faba is exceedingly large (ca. 13 Gb) provides a valuable tool for future transcript profiling and also enriches the genetic resources available for this important legume crop species (Arun-Chinnappa and McCurdy, 2015). Development of bacterial artificial chromosome (BAC) library in White clover, sought a large-scale BAC-end sequencing strategy has the potential to anchor a significant proportion of the genome of white clover onto the gene-space sequence of M. truncatula (Febrer et al., 2007). The first draft of the genome sequence of pigeonpea, and development of markers will be useful for fingerprinting and diversity analysis of pigeonpea germplasm and molecular breeding applications (Dubey et al., 2011; Singh et al., 2012). In 2012, Illumina next-generation sequencing platform to generate 237.2 Gb of sequence applied in pigeonpea (Varshney et al., 2012).

In the whole genome sequencing, genetic map provides an essential framework for accurate and efficient genome assembly and validation. The analysis revealed an overall conserved marker order, although some localized inversions between draft genome assembly and the genetic map were detected (Deokar et al., 2014). Many vectors, especially artificial chromosome vectors, have been developed for genome-scale mapping and sequencing. These data showed that this genomic library was reliable for further molecular research in Lotus japonicus (Guo Xi-ZhiGuo et al., 2004).

**Comparative Genomics**

Within the legume plant family, glimpses of synteny have also been observed. Characterizing syntenic relationships in legumes is important in transferring knowledge from model legumes to crops that are important sources of protein, fixed nitrogen, and health-promoting compounds (Mudge et al., 2005). The Lotus and Medicago genomes share a minimum of 10 large-scale synteny blocks, each with substantial collinearity and frequently extending the length of whole chromosome arms (Cannon et al., 2006). 56 conserved synteny blocks were identified between eggplant and tomato (Hirakawa et al., 2014). The December 22 2011 issue of Nature published the result of syntenic relationships between Medicago, Glycine, Lotus and Vitis illustrated in Circos diagram (Young et al., 2011) (Figure 1). Phylogenetic interpretation of DNA sequence data from multiple genomic regions has become the gold standard for species delimitation and population genetics (Kohn, 2004).

The interaction between genomics and quantitative genetics has been a two-way street. Genomics contributed genetic markers and genetic maps making it possible to study quantitative trait loci (QTLs), and quantitative genetics contributed new theories and computational techniques to deal with the data generated by QTL studies (van Buitjelenen, 2001). It will be possible to integrate-based genomic tools into approaches as genetic approaches aiming at developing efficient marker-assisted selection strategies for the development of resistant varieties for example, or for designing models for predicting the behaviour of the plant in the field for greater resistance (Adam-Blondon, 2007).

Gene discovery holds enormous promise, but depends heavily on comparative genomics, capitalizing on genomic information (Burdon and Wilcox, 2007). Comparative genomics approach are often used to investigate crops with narrow genetic base of cultivars, and to evaluate the
evolution of a complex gene cluster that is associated with several disease resistance (R) genes of known function (Ashfield et al., 2012; Bertioli et al., 2003; Sharma et al., 2013) Comparative genomics provides a tool to utilize the exponentially increasing sequence information from model plants to clone agronomically important genes from less studied crop species (Huang et al., 2005). Comparative genomics also plays a very important role in identification of genes/QTLs linked to the resistance genes (Babu et al., 2014). With access to sequences various model organisms and many important microbial pathogens, structural biology is on the verge of a dramatic transformation. Structural genomics will yield a large number of experimental protein structures (tens of thousands) and an even larger number of calculated comparative protein structure models (millions) (Burley, 2000).

The emergence of "omic" technologies and the establishment of model legume plants, such as Medicago truncatula and Lotus japonicus are promising strategies for understanding the molecular genetic basis of stress resistance (Dita et al., 2006). The legume species M. truncatula is a model plant to help in the identification of agronomically important genes in crop legumes, e.g. disease resistance (Journet et al., 2001).

QTL mapping and synteny analysis identified genomic loci harboring resistance factors has offered considerable promise and experimental power with varying degrees of success (Colwell, 2002; Muchero et al., 2011). The potential of GWAS approaches likely bearing resistance alleles of significant effect useful in breeding programs (Debibakas et al., 2014). Genome-wide analyses of gene function and gene expression are beginning to yield valuable information in many areas of biological research, and these genomic tools are now being applied to crop pest and disease research (Keon et al., 2003), thus strategy of applying a sequence-to-phenotype (or functional genomics) paradigm applied to the discovery of R-genes (Kamoun et al., 2002).

**Figure 1.** Homologous gene pairs were identified for all pairwise comparisons between M. truncatula (Mt), G. max (Gm), L. japonicus (Lj) and V. vinifera (Vv) genomes, syntenic regions are associated with the ancestral whole-genome duplication (WGD) occurred approximately 58 Myr ago (Young et al., 2011).
Conclusions
Current endeavours to find disease-causing genes have received a growing appreciation of the extent of conservation of genes, the extent to which syntenic will be a useful tool throughout all legume species. The impact of the discovery of conserved synteny is also most important in the consideration of genomics programs for the major legume genomes. Genomics, through deciphering allelic effects on phenotypes and identifying patterns of adaptive variation at the landscape level, will in the future constitute a useful tool, if cost-effective, to design conservation strategies for plants (Gonzalez-Martinez et al., 2006), including disease resistance breeding program. However, the discovery still needs to be validated in a much larger sample size, where a high hurdle that many putative biomarkers fail to make. Poste (Poste, 2011) explained that although over 150,000 papers claimed to have found disease biomarkers in human, only 100 biomarkers were routinely used in clinic. Perhaps the same case may occurred in legumes, and future research on linking biomarkers comparative omics will help to fulfill this challenge.

References


cowpea (Vigna unguiculata L.) with the root parasitic angiosperm Striga gesnerioides. Plant Methods 8.


Other Publications of Interest