‘Omics’ approach to identify factors involved in Brassica disease resistance

Marta Francisco¹, Pilar Soengas¹, Pablo Velasco¹, Vijai Bhadauria², Maria E. Cartea¹ & Victor M. Rodríguez¹*

¹ Group of Genetics, Breeding and Biochemistry of Brassicas. Misión Biológica de Galicia (MBG-CSIC), Apartado 28, 36080 Pontevedra, Spain

² Crop Development Centre. Department of Plant Sciences, 51 Campus Drive. University of Saskatchewan, Saskatoon, SK S7N 5A8 Canada

* Corresponding author: Victor M. Rodriguez; Tel: +34 986 85 48 00; FAX: +34 986 84 13 62; e-mail: vmrodriguez@mbg.csic.es
Abstract

Understanding plant’s defense mechanisms and their response to biotic stresses is of fundamental meaning for the development of resistant crop varieties and more productive agriculture. The Brassica genus involves a large variety of economically important species and cultivars used as vegetable source, oilseeds, forage and ornamental. Damage caused by pathogens attack affects negatively various aspects of plant growth, development, and crop productivity. Over the last few decades, advances in plant physiology, genetics, and molecular biology have greatly improved our understanding of plant responses to biotic stress conditions. In this regard, various 'omics' technologies enable qualitative and quantitative monitoring of the abundance of various biological molecules in a high-throughput manner, and thus allow determination of their variation between different biological states on a genomic scale. In this review, we have described advances in ‘omic’ tools (genomics, transcriptomics, proteomics and metabolomics) in the view of conventional and modern approaches being used to elucidate the molecular mechanisms that underlie Brassica disease resistance.
Introduction

The family Brassicaceae (syn. Cruciferae) is one of the crucial plant families for humans and animals and supplies several products from various plant parts. Since ancient times, Brassica crops have been used for many purposes, including vegetables, oilseeds, feed, condiments, fodder, green manure and even medical treatments. The genus Brassica includes a group of six interrelated species of worldwide economic importance. The principal vegetable species is B. oleracea, which includes vegetable and forage crops, such as kale, cabbage, broccoli, Brussels sprouts, cauliflower, and others; B. rapa includes vegetable forms, such as turnip, Chinese cabbage, and Pak Choi, along with forage and oilseed types; B. napus crops are mainly used as oilseed (rapeseed), although forage and vegetable types like leaf rape and “nabicol” are also included. Finally, the mustard group which is formed by three species, B. carinata, B. nigra and B. juncea, is mainly used as a condiment because of their seeds although leaves of B. juncea are also consumed as vegetables in Asian countries.

The productivity and quality of these crops are seriously affected by several diseases, which result in substantial economic losses every year worldwide. The increase in the economic importance of Brassica crops has led to an increase in research on the plant host responses to pathogens and the mechanisms underlying the resistance. Several diseases are known to occur on Brassica crops affecting yield and quality in the field and in storage (Table 1). Among them, black rot and clubroot could be considered the most important diseases. Black rot is caused by the bacterium Xanthomonas campestris pv. campestris (Pammel) Dowson and is considered to be one of the most devastating soilborne disease in Brassica crops worldwide (Lema et al., 2012; Vicente et al., 2001). The disease is favored by warm, humid conditions and can spread rapidly
from rain dispersal and irrigation water. Typical disease symptoms include V-shaped yellow lesions starting from the leaf margins and blackening of the veins (Vicente et al., 2006; Vicente and Holub, 2013). Considerable research into the management of black rot in vegetable Brassica crops has been done, but the majority of commercial cultivars of B. oleracea and B. rapa are highly susceptible (Lema et al., 2012). Clubroot is caused by the protist Plasmodiophora brassicae which infects the majority of cruciferous and it is particularly important in horticultural Brassica production. P. brassicae causes galls on roots of infected plants which interfere with the plant uptake of water and nutrients. Under warm conditions, plants often wilt which is usually the first sign of clubroot infection (Rennie et al., 2015). Ideal conditions for clubroot infection are acid soils, high soil moisture, warm temperatures and the presence of a susceptible brassica host (Strehlow et al., 2014)

Besides Xanthomonas, other bacteria could cause diseases in Brassica spp. For instance, Erwinia carotovora and Pseudomonas spp. are common pathogens to most Brassica crops and cause a soft mushy breakdown (Table 1). P. syringae pv. maculicola (McCulloch) is specially virulent. The bacterium can be seed-borne and survive on plants and crop debris and infects early vegetative plant stages. Development of infection is favored by prolonged cool, wet conditions and can develop rapidly following plant stress.

Fungi could also cause important diseases in Brassica spp. Alternaria spp. cause leaf spot that affects the quality of the products and can become a major problem during storage (Lee and Hong, 2015). Most of the important economic Brassica crops are hosts of A. brassicae (Table 1). This fungus produces small dark spots on older leaves in cool wet conditions, which are surrounded by a ring of lighter tissue. The disease can be carried over on crop residues and weeds. Other fungi such as Leptosphaeria maculans
or *Sclerotinia sclerotiorum* (Lib.) de Bary attack a wide range of *Brassica* crops (Table 1). The former fungus causes dry rot/blackleg that commonly affects the stems and leaves of vegetable brassicas. Numerous vegetable and oilseed *Brassica* crops are hosts of this disease (Table 1). Stem damage can lead to cankering and severing of the plant at the base. *S. sclerotiorum* causes the so-called stem rot disease. The disease is mainly on aboveground parts, producing a cottony white mold. Finally, the mold is replaced by numerous hardened sclerotia that gradually change from white to black (Saharan and Mehta, 2008).

There are many viruses that affect *Brassica* crops. Symptoms vary depending on the plant host, age, variety, weather conditions and nutritional status. Viruses often have a number of alternative *Brassica* hosts and are usually spread from plant to plant by insects (e.g. aphids, thrips) or fungal vectors. The three main *Brassica* viruses are beet western yellows virus (BWYV), cauliflower mosaic virus (CuMV) and turnip mosaic virus (TuMV) (Table 1). Infected plants show patterns, distortion and mottling in the leaves and often die if infection occurs early (Hunter et al., 2002). The application of aphicides can prevent their introduction; however, correct timing is difficult to achieve. Seed treatments can give seedlings up to some weeks protection, often long enough to protect through the critical period when aphids are flying (McSorley and Frederick, 1995; Raybould et al., 1999).

**‘Omic’ studies in *Brassica*-pathogen interactions**

Plant infection by pathogenic microbes involves many dynamic changes in molecular communication and adaptation of host and pathogen physiology. Significant advances have been made in the past few decades in understanding the defense mechanisms
associated with resistance in Brassica crops, and to date many genes governing resistance have been identified and used in crop improvement. However due to the complexity of the genetic and molecular processes implicated in defense responses they need to be researched more extensively. In this regard, the rapid advances in ‘omic’ technologies provide an opportunity to generate new information either at the genomic, transcriptomic, proteomic or metabolomic levels. In the following sections, we provide an overview of integrated high-throughput genotyping technologies and functional ‘omic’ tools recently used for the identification of new genes, metabolic pathways and proteins involved in plant-pathogen resistance of Brassica crops.

Genomics

Genomics is a discipline in genetics that applies recombinant DNA, DNA sequencing methods, and bioinformatics to sequence, assemble, and analyze the function and structure of genomes. The field includes efforts to determine the entire DNA sequence of organisms and fine-scale genetic mapping. The first species of Brassicaceae being sequenced was Arabidopsis thaliana (Kaul et al., 2000). Genomes of the most important Brassica crops from the economical point of view have been recently sequenced and assembled combining Sanger and next-generation sequencing strategies (NGS). Genome sequences included those of Chinese cabbage (B. rapa), cabbage and kale (B. oleracea) and rapeseed (B. napus) (Chalhoub et al., 2014; Liu et al., 2014; Parkin et al., 2014; Wang et al., 2011). The genomes of B. rapa (A) and B. oleracea (C) have suffered a whole genome triplication (WGT), containing each three sub-genomes (Liu et al., 2014; Parkin et al., 2014; Wang et al., 2011), being this a characteristic of the Brassicaceae tribe (Li et al., 2014). While the genomes of B. oleracea and B. rapa are
highly similar in terms of total gene clusters/sequences and the gene number in each
cluster, there are also a large number of species-specific genes in the two species. *B.
napus* is an amphidiploid between *B. rapa* and *B. oleracea* (AC genome). Most
orthologous gene pairs in *B. rapa* and *B. oleracea* remain as homologous in *B. napus*
(Chalhoub et al., 2014). Genomes sequencing projects for *Brassica* crops have produced
vast amounts of sequence data that will provide useful information for genetic studies
related to disease resistance. Sequences of the A, C and AC genomes, their assembly
and predicted genes and associated annotations are available in several web pages which
are summarized in: [http://www.brassica.info/resource/genome.php](http://www.brassica.info/resource/genome.php). The availability of
these reference genomes enhances our understanding of the genome architecture and the
evolution of *Brassica* species, as well as facilitates identification of genes associated
with important traits for breeding (Lee et al., 2015).

Besides, the genomes of other species from the *Brassicaceae* family closely
related to *Brassica* crops, have been recently sequenced as those of *Eutrema
salsugineum, Capsella rubella, Leavenworthia alabanica, Sisymbrium inio, Aethionema
arabicum* or *Arabidopsis lyrata* (Hu et al., 2011; Slotte et al., 2013; Yang et al., 2013).

The availability of genome sequences of *Brassicaceae* species allows carrying
on comparative genomics, to study basic biological similarities and differences as well
as evolutionary relationships between organisms. The close phylogenetic relationship
between the *Brassica* species and the model plant *A. thaliana* makes the transfer of
knowledge from *A. thaliana* to *Brassica* crop improvement straightforward (Cho et al.,
2015). Several studies have established chromosomal collinearity between *Brassica* and
*A. thaliana*. The QTL of resistance to clubroot Crr1 is located on a region called A8 on
*B. rapa* (Suwabe et al., 2012). After fine mapping, the region on A8 corresponded to a
disease resistance gene cluster in the *A. thaliana* genome. The location and order on the
markers showed good correspondence with those in *A. thaliana* with a few rearrangements such as inversions and insertions. This genomic region of *Brassica* is conserved and has a potential as a disease resistance gene cluster.

Moreover, the availability of whole genome sequences has led to the identification and validation of novel disease-resistance alleles (Hayward et al., 2012). Plant disease resistance (R) genes which specifically interact/recognize with corresponding pathogen avirulence (avr) genes are considered as plant genetic factors of a major layer. The interactions of this gene-for-gene manner activate the signal transduction cascades that turn on complex defense responses against pathogen attack and this is called incompatible interaction (Yu et al., 2014). Whole genome analysis of R gene families within cultivated and wild varieties of a host species, made available through genome sequencing, may not only enable the identification of novel genes, but extensive phylogenetic and evolutionary analyses may also infer the nature of R and Avr gene product interactions (Hayward et al., 2012). Most of the R genes in plant kingdom are members of NBS-LRR (nucleotide-binding site-leucine rich repeat) proteins. Recent whole genome sequence data enabled the genome wide identification, mapping and characterization of candidate 159 NBS-containing R genes in *A. thaliana* (Meyers et al., 2003). The genome sequence of *B. rapa* was used to identify 92 non-redundant NBS-encoding in approximately 100 Mbp of *B. rapa* genome sequence (Mun et al., 2009). More recently, 157 and 206 NBS-encoding genes in *B. oleracea* and *B. rapa* were identified, respectively (Yu et al., 2014). NBS-encoding genes exhibited differential expression pattern in different tissues and several genes were induced by wounding. Despite the fact that *B. rapa* has a significantly larger genome than *A. thaliana*, *B. rapa* contains relatively small number of NBS-encoding genes, presumably
Transcription factors (TF) play pivotal functions in signal transduction to activate or suppress defense response genes and regulate the interactions between different signaling pathways. The AP2/ERF superfamily is one of the largest groups of TFs in plants. Extensive research has confirmed that AP2/ERF TFs are involved in plant growth and development, hormone response, and biotic or abiotic stress responses (Liu et al., 2013). ERF genes could improve plant resistance because their overexpression enhanced resistance to various diseases and improved tolerance to drought, salt, and freezing in transgenic plants (Liu et al., 2013). The *Brassica* Database was surveyed to gain further information on the AP2/ERF superfamily and its subclade CRFs in Chinese cabbage. A total of 281 members were identified in this superfamily. Up to 35 tandem duplicated genes and 252 segmental duplicated genes were found among the 281 AP2/ERF TFs, while 30 tandem duplicated genes and only 75 segmental duplicated genes were found in *A. thaliana*, suggesting that the expansion of BrAP2/ERF TFs after speciation from *A. thaliana* is mainly attributed to segmental duplication events during the WGT.

With the availability of the *Brassica* A (Wang et al., 2011), C (Liu et al., 2014; Parkin et al., 2014) and AC genome (Chalhoub et al., 2014) sequences it is possible to identify high-density, genome-wide SNPs in *Brassica* spp. (Dalton-Morgan et al., 2014). High-density SNP arrays can be a cost-effective alternative for genome-wide polymorphism screens for genome-wide association (GWAS) studies of resistance to diseases, but also by breeders as a tool for comprehensive genome-wide screens of elite germplasm and breeding pools. Genome-wide high-density array in *B. napus* has been described recently (Dalton-Morgan et al., 2014) and it has been demonstrated its utility...
in genotyping R resistance genes to black leg on chromosome A07. A total of 21,311 SNP markers were developed by Park et al. (2010) by re-sequencing 1,398 STSs in eight genotypes of Chinese cabbage (B. rapa). 141 SNPs related to disease resistance and leaf traits were selected (Ahn et al., 2014). Among them, 20 SNP primers were polymorphic in a DH mapping population differing in resistance to clubroot, turnip mosaic virus and soft rot and they could be further used for the detection of QTLs and fine mapping studies of disease resistance of Chinese cabbage. From the same original set of SNPs markers designed by Park et al. (2010), 693 SNPs were tested for amplification in the B. oleracea genome, from this 425 were successfully applied into B. oleracea, suggesting that it is possible to apply SNPs markers developed base on the B. rapa genome to B. oleracea (Cho et al., 2015).

The application of GWAS to plants has greatly increased the resolution of QTL detection. Fopa Fomeju et al. (2014) by using a GWAS analysis identified genomic regions of B. napus linked to the resistance to stem canker, caused by the fungal pathogen L. maculans. A panel of 116 oilseed rape spring varieties was genotyped with 3228 SNPs. 321 markers were linked to the resistance, corresponding to 64 genomic regions. These genomic regions were relatively equally distributed on the A (53%) and C (47%) genomes of B. napus. 44% of these regions are duplicated homologous regions, which suggests structural and functional conservation of genetic factors involved in this trait in B. napus.

Efforts to understand crop genomes have been boosted in recent years by developments in NGS technologies. Dramatic cost reductions accompanying the rapid increase in scale of these systems today make DNA sequencing an accessible and powerful option for high-throughput analysis (Edwards et al., 2013). Re-sequencing is
used to identify genetic variation between individuals, which can provide molecular
genetic markers and insight into gene function (Tollenaere et al., 2012).

Using the sequencing-by-synthesis method, a total of 7.0 Gb Illumina pair-end
reads were generated from each one of the parental lines of a mapping population of
cabbage (Lv et al., 2013) with different levels of resistance to *Fusarium* wilt (caused by
the fungus *Fusarium oxysporum*). InDel markers were designed according to the
reference genome sequence of cabbage and the whole-genome resequencing of the two
parents. A resistance gene (FOC1) was located on C06 flanked by two InDel markers.
Afterwards Lv et al. (2014) designed additional InDel primers in that region to map the
FOC1 gene more precisely. Two TIR-NBS-LRR type R genes were considered as
candidates genes underlying the variation for *Fusarium* wilt resistance.

Black rot is one of the most devastating diseases to crucifers including *B.
oleracea*. Up to 14 QTLs were described as linked to resistance in cabbage (Lee et al.,
2015). Two parental lines of a mapping population of cabbage were re-sequenced up to
20x genome coverage and a genome-wide survey for SNPs was conducted.
Approximately, 1.20 millions of SNPs were identified in the parents. Polymorphic
SNPs between both parents were used to design dCAPS markers for those regions with
low marker density. Finally, 103 markers were added to develop a higher density map.
NBS-LRR genes within the QTLs regions were searched. Four QTL regions contained
21 candidate R genes. Two candidate disease resistance genes in BRQTL-C1 and seven
genes in BRQTL-C3 were found as gene clusters (Lee et al., 2015).

Little information is available on the genetic control of quantitative resistance to
blackleg. A QTL for blackleg resistance was mapped in a DH population of rapeseed
(*B. napus*) and corresponds to the R gene Rlm4 placed on A07. This region was aligned
with the reference genome sequence of *B. rapa* on which 18 candidate R genes were
identified (Tollenaere et al., 2012). Parents of the DH population were re-sequenced (read coverage of 10x in both). Sequences were aligned with the reference genome of *B. rapa* and the 18 candidate genes were characterized in the parental lines following Illumina paired-end sequencing of the parents. Two candidate genes (BLR2 and BLR11) were selected based on the sequence identity of known resistance genes and sequence divergence in the NGS reads of both parents.

**Transcriptomics**

To effectively combat pathogen invasion, plants have developed sophisticated defense strategies that initiate with an extensive transcriptome reorganization that presages changes in plant metabolism. When the infection plays out, the plant metabolism often represents a shifting mixture of disease resistance responses and disease susceptibility responses. Over the last decades, studies on the transcriptome mechanisms that underlie the effects of plant pathogenic infection have received a great deal of attention. At a staggering pace the development of new technologies to perform genome-wide transcriptomic analyses (microarray based expression profiling methods, the availability of genomic and/or EST (expressed sequence tag) sequence data for some plant species and the recent development of the NGS) has allowed significant progress in the characterization of the transcriptomic plant-pathogen related response.

Speaking strictly the transcriptome is defined as the total set of transcripts (including mRNA, tRNA, rRNA and small non-coding RNAs) present in a given cell, tissue or organism. Techniques to analyze the transcriptome are mainly focused on the relative or absolute quantification of mRNA. These techniques could be divided in two major groups, those based on hybridization and those based on amplification.
Microarrays are the paradigmatic example of the former group and were the first platform developed to study genome-wide expression. DNA microarrays provide a simple vehicle for exploring the transcriptome in a way that is both systematic and comprehensive (Brown and Botstein, 1999). This platform is based on the immobilization of discrete DNA sequences (probes) in an array which are used to interrogate the mRNA of a sample (target). Probe construction is a crucial step on microarray design. Regardless its chemical nature, probes have to hybridize very specific sequences and thus their design requires a basic knowledge of the genomic sequence of the species. For this reason the use of microarray technology has been limited to model organisms for years (Witzel et al., 2015). In the particular case of Brassica spp. early transcriptomic studies took advantage of the high degree of synteny in gene sequences between A. thaliana and Brassica spp. This highly genetic similitude has allowed the use of the A. thaliana microarrays to examine transcriptional changes associated with the response of B. napus cultivars to S. sclerotiorum infection at different time points following pathogen infection (Liu et al., 2005; Yang et al., 2007; Zhao et al., 2007) or to analyze the transcriptomic response to TuMV (turnip mosaic virus) infection in B. rapa (Li et al., 2014).

Nowadays, advances in the acquisition of Brassica spp. genome sequences have prompted to the development of specific microarrays. Two major groups of microarrays are available to perform genome-wide transcriptome studies in Brassica spp.; spotted microarrays and oligonucleotide microarrays being both alternatives successfully used to characterize the Brassica response to different pathogens. Narusaka et al. (2006) used a spotted microarray with 1820 cDNA clones selected from 20166 non-redundant sequences of cDNA libraries of Chinese cabbage to investigate the transcriptome induced changes in tissues inoculated with the fungal pathogen Colletotrichum...
higginsianum. A custom-made cDNA microarray with 26,090 oligonucleotides was also employed by Zhao et al. (2009) to monitor gene expression in *B. napus* in response to *S. sclerotiorum* insult. A different strategy was followed by Yin et al. (2006), who treated *B. napus* plants with oligochitosan to elicit the plant immune response and characterize this response using a microarray containing 8,095 ESTs. A collection of *Brassica* specific oligonucleotide microarrays are nowadays commercially available (Love et al., 2010; Trick et al., 2009), but so far these have been seldom used in *Brassica*-pathogen interactions. Just recently, Jung et al. (2013) studied the transcriptional profiling of a transgenic Chinese cabbage using the *B. rapa* 24K oligo array challenge with *Pectobacterium carotovorum* subsp. *carotovorum* (Pcc).

The alternative to microarray platforms relies on technologies based on amplification. The recent acquisition of *Brassica* species genome sequences together with the advance in NGS, have allowed accurate measurement of expression levels of genes as well as detection of allele specific expression and alternative splicing variants. Two protocols have been developed to quantify level of expressions: tag-based sequencing approach (Digital Gene Expression (DGE)) and RNA sequencing (RNA-seq). DGE includes different techniques such as serial analysis of gene expression (SAGE) (Yamamoto et al., 2001), SuperSAGE (Matsumura et al., 2010) or cap analysis of gene expression (CAGE) (Kodzius et al., 2006). All these techniques are based on the generation of short cDNA tags from an mRNA population in several steps. More straightforward protocols are used for RNA-Seq. A typical RNA-Seq experiment starts with mRNA that is subsequently converted into cDNA to form an RNA-Seq library. By sequencing the millions of DNA fragments in the library (known as ‘reads’) with NGS, an accurate measure of the relative abundance of each transcript and splice variants can
be obtained. In principle, any high-throughput sequencing technology could be used in both approaches.

Large-scale and genome-wide gene expression profiling methods have been applied to analyze *Brassica* crops infected by pathogens. Expression patterns of genetically transformed rapeseed plants with *hrf2*, a gene conferring rapeseed resistance to *S. sclerotiorum*, were compared with untransformed genotype, observing the breadth of molecular mechanism of hairpin-mediated plant responses related to defense, plant-hormone biosynthesis, catabolism, and signal transduction (Wang et al., 2015). RNA-Seq has been employed recently to elucidate resistance mechanisms involved in plant-pathogen interactions including fungi, bacteria and virus (Chu et al., 2014; Kalischuk et al., 2015; Kim et al., 2015; Lloyd et al., 2014; Shimizu et al., 2014).

Compared to other plants little information is available about the genome-wide expression response of *Brassica* spp. to pathogen insult. Special attention has been paid to the response of these species to fungal attack. Experiments on transcriptome profiling analysis of *Brassica* cultivars infected with fungal pathogens, led to the identification of more than 300 differentially expressed genes at least two-fold as compared to uninoculated controls and a large part of these genes exhibits temporal and quantitative differences between resistant genotypes and susceptible ones (Kim et al., 2015; Liu et al., 2005; Lowe et al., 2014; Narusaka et al., 2006; Shimizu et al., 2014; Yang et al., 2007; Zhao et al., 2007). Expressed genes included those encoding defense-associated proteins, transcription factors belonging to the zinc finger (WRKY, APETALA2 and MYB classes), phytohormone-responsive genes (jasmonic acid, ethylene and auxin synthesis enzymes) and cell wall structure genes. In addition, changes in the expression of genes encoding enzymes involved in carbohydrate and energy metabolism appeared to be directed towards shuttling carbon reserves to the tricarboxylic acid cycle and
generating reactive oxygen species (Kim et al., 2015; Lowe et al., 2014; Shimizu et al., 2014; Zhao et al., 2009; Zhao et al., 2007). Transcripts from genes encoding enzymes involved in glucosinolate and phenylpropanoid biosynthesis were highly elevated after fungus infection, suggesting that secondary metabolites are also components of the defense response in *Brassica* spp. (Shimizu et al., 2014; Zhao et al., 2009). The integration of metabolomics with transcriptomic and genomic platforms has frequently been used as a strategy to identify candidate genes involved in the regulation of the levels of specific metabolites in plant systems. Thus, some of the candidate defense genes detected by transcriptomic approaches were also identified by integration with previously mapped QTLs (Zhao et al., 2007), and validated by qRT-PCR and Northern blot analysis (Chu et al., 2014; Liu et al., 2005; Yang et al., 2007; Zhao et al., 2009).

Transcriptomic response to bacterial or virus pathogens has been extensively studied in *A. thaliana*, but not so in *Brassica* spp. Efforts to obtain *B. rapa* genotypes resistant to *P. carotovorum* prompted the development of transgenic lines expressing a human cathelicidin antimicrobial peptide (Jung, 2013). The human peptide confers a moderate resistance to the pathogen by activating the expression of defensins, resistance-responsive proteins and LTP (lipid transfer proteins). Bacterial and virus pathogens have also been used to activate resistance gene expression of *Brassica* spp. (Kalischuk et al., 2015; Sarosh et al., 2009). Immunization of *B. napus* plants with *Bacillus amyloliquefaciens* prevents the attack of *Botrytis cinerea* (Sarosh et al., 2009). Interestingly, in spite of the root treatment of plants with *B. amyloliquefaciens* more transcripts were activated in leaves than in roots indicating a systemic effect in priming. Likewise, low dose of purified dsCaMV virus induce a complex transcriptome reactivation that allows a transgenerational resistance to pathogens. Though the
mechanism involved in this response remain elusive, transcriptomic analysis point to epigenetic mechanisms (Kalischuk et al., 2015).

A special case of a transcriptomic study is the expression profile of small RNAs (including miRNAs) in response to pathogen challenge. Small RNAs are noncoding RNAs that play an important regulatory role in the growth and development of eukaryotes. In plants they are known to regulate the expression of a number of key developmental and stress-related genes and plant innate immune receptors (Eckardt, 2012). During the last years technological advances have allowed the study of the genome-wide expression profile of sRNAs. Shen et al. (2014) used this approach to analyzed the role of miRNAs in the oilseed rape-Verticillium longisporum interaction using two RNA libraries made from V. longisporum infected/non-infected oilseed rape roots and employing B. rapa and B. oleracea genomes as references for miRNA prediction and characterization. A high throughput analysis identified two classes of sRNAs of B. campestris that modified their expression after infection with E. carotovora, host 28-nt sRNAs and sRNA homologous to the pathogen genome (Sun, 2014). Interaction with the protist P. brassicae also activates the expression of sRNAs in B. napus (Verma et al., 2014). This sRNAs are involved in the regulation of the expression of several transcription factors related to plant resistance.

Proteomics

Given that the language of plant-pathogen interactions is largely in proteins and is the final executors of most biological processes, proteomics is an important tool in translational genomics as it can translate plethora of genomic information into functional information. Proteomics is the large-scale functional analysis of proteins
extracted from intact organisms, tissues, individual cells, or cell compartments, at
defined time points during development or under specific conditions. Comparative
proteome analysis, could lead to a more comprehensive understanding of biotic stresses
in economical important *Brassica* crops as it can reveal factors involved in resistance
against various diseases as proteins are directly related to functions. These factors can
then be used in developing molecular markers to introgress resistance into cultivars
through marker-assisted selection breeding or genome editing.

Many initial proteomic approaches were based two-dimensional gel
electrophoresis (2-DE), a good choice for rapidly identifying major proteome
differences in healthy versus inoculated plants. However it suffers from some ongoing
concerns regarding quantitative reproducibility and limitations on the ability to study
certain classes of proteins. Therefore, in recent years attention has been paid on the
development of alternative approaches, such as promising gel-free proteomics. With the
appearance of mass spectra (MS)-based proteomics, an entirely new toolbox has
become available for quantitative analysis. Although these novel approaches were
initially pitched as replacements for gel-based methods, they should probably be
regarded as complements to rather than replacements of 2-DE. Protein spots selected
from the gels are picked, trypsin digested, and subjected to ionization by ESI or matrix-
assisted laser desorption/ionization (MALDI), both coupled to MS analyses.
Subsequently, the peptide masses are used to query peptide mass databases in order to
identify the proteins (Rabilloud et al., 2010). This proteomics approach has been
successfully applied to some host-pathogen interaction studies in *Brassica* vegetables.

Canola (*B. napus*) is one of the important oil-seed crops worldwide, and clubroot
and blackleg are the two main biological constraints to canola production. Clubroot
caused by an obligate biotrophic protist *P. brassicae* is creating havoc in canola
growing regions. Source of resistance has been identified in European fodder turnip B. *rapa* ssp. *rapifera* and successfully introgressed into oilseed rape (*B. napus*) (Bradshaw et al., 1997; Diederichsen et al., 2009; Hirai, 2006; Piao et al., 2009; Yoshikawa, 1981). This genetic resistance coupled with crop rotation can break the prolonged survival cycle of this pathogen in soil (up to 20 years). Kaido et al. (2007) carried out a comparative proteome analysis between a pair of resistant and susceptible turnip cultivars to the causal agent of culbrot disease. They achieved 2-DE to detect differences in protein profiles between resistant and susceptible cultured roots in the early responses against the given *P. brassicae* resting spores. These proteins, especially the up-regulated ones in resistant roots, must include specific agents involved in the resistance to clubroot infection.

Time-course proteomics revealed major changes in canola metabolism leading to susceptibility to *P. brassicae* (Cao et al., 2008). The authors used 2-DE coupled ESI-MS to identify 20 proteins in canola roots differentially regulated upon *P. brassicae* infection. Among these, there can be found reactive oxygen species-detoxification enzymes such as Cu/Zn sodium dismutase and cytochrome c oxidase, proteins implicated in metabolic pathway, enzymes such as S-adenosylmethionine synthetase, adenosine kinase, triose phosphate isomerase and Glycine-rich RNA-binding protein and the lignin biosynthesis enzyme caffeoyl-CoA O-methyltransferase. Blackleg is another major disease of oilseed rape worldwide. Twenty-four differentially expressed proteins in *B. napus* resistant cultivar Surpass 400 upon challenge with virulent and avirulent isolates of *L. maculans* were identified by the 2-DE coupled with matrix-assisted laser desorption ionization time-of-flight mass spectrometry (TOF-MS devices) (Marra et al., 2010). Among these there can be found stress- and defense associated proteins, such as heat shock protein, peptidyl-prolyl cis-trans isomerase, ribosome
inactivating protein, ascorbate peroxidase and disulphide oxidoreductase. Functional validation of clubroot and blackleg resistance associated proteins can be achieved by generating transgenic plants using genome editing tools, such as Zn finger nuclease, TALENs and CRISPR/Cas9. Sun et al. (2014) studied proteome changes in the seedling leaves of the non-heading Chinese cabbage at different time points following inoculation with *H. parasitica* using 2-DE in combination with matrix-assisted laser desorption/ionisation TOF-MS. They found 91 protein spots related to the resistance response. These proteins were assigned to different functional categories, such as amino acid and carbohydrate metabolism, photosynthesis and photorespiration, protein metabolism, signal transduction, redox homeostasis, and ethylene biosynthesis.

Recently developed shotgun proteomics approaches, such as iTRAQ (Isobaric tags for relative and absolute quantitation) have broader implication in plant disease resistance as these approaches allow multiplexing of samples and therefore, simultaneously compare and quantify differentially regulated proteins in susceptible and resistant genotypes. These technical advancements coupled to well-designed experiments will significantly reveal the protein function in plant stress- and defense and provide a wealth of information on plant proteome changes occurring in response to biotic stresses.

**Metabolomics**

For numerous organisms, complete genomes have been sequenced and transcriptome and proteome studies have been described. Only recently, metabolome analysis using MS platforms attracted attention (Hill and Roessner, 2013). Metabolomics is the analysis of the small molecules (metabolome) of an organism in a determined moment.
This analysis provides the phenotypical response at the metabolic level under a particular environmental condition or stress. It is considered a very ambitious field due to the high number of chemically distinct molecules present in a typical plant sample, estimated to be at least 100,000 (Witzel et al., 2015). As stated by Sardans et al. (2011), metabolomics provides a better analysis of the different response capacities conferred by the phenotypic plasticity of each species, allowing ascertaining what metabolic pathways are involved in a phenotypic response. The metabolome consists of two types of compounds, the primary metabolites, compounds shared by different organisms, involved in the basic functions of the living cell, such as respiration, and the secondary metabolites. The later are species specific and play a role in the interaction with the environment, e.g. plant defense against pests and diseases (Verpoorte et al., 2007).

The analysis of metabolites is done by analytical platforms which combine several techniques like nuclear magnetic resonance spectroscopy (NMR), MS, gas chromatography (GC), liquid chromatography (LC), and capillary electrophoresis (CE). As explained by Witzel et al. (2015) changes in the primary and secondary metabolite pool have been characterized using a GC-MS approach, while CE-MS is more suitable for characterizing the products associated with the central metabolism. NMR is appropriate for the targeting of phenolic compounds, carbohydrates, organic acids and amino acids; ESI is used for the analysis of semi-polar metabolites. Together, these technologies can identify and quantify a wide range of primary and secondary metabolites. Besides, it is possible to add high mass resolution, MS/MS fragmentation patterns and UV spectra by the use of TOF-MS devices. Recent software developments have improved the capacity to recognize different metabolites. Profiles based on mass, retention time and signal amplitude provides the data required for filtering biomarkers. Typically a data processing pipeline can be divided into two steps: data processing...
(filtering, feature detection, alignment, and normalization) and data analysis (algorithm selection, training, evaluation, and model examination) (Witzel et al., 2015). To spread the use of metabolomics it is necessary to have public databases to compare the obtained results in the lab. There are two important characteristics to have a good metabolome analysis: the reproducibility to study the effect of different conditions in a concrete plant and the ease of quantitation and identification the number of metabolites to be measured (Verpoorte et al., 2007). Based on those characteristics Veerporte et al. (2007) established three different groups of methods: 1- Chromatographic methods, 2- MS and 3- NMR, that can be reduce to two methods, the combination of 1 with 2 or 3. MS is now the most extended method but has a problem with the reproducibility due to the previous ionization step and the different mass spectrometers. NMR is a physical measurement of the resonances of the magnetic nuclei in a strong magnetic field. Each compound has a highly specific spectrum. This technique is highly reproducible which makes NMR the most suited method for a public domain metabolomics database. Chromatographic methods have shown to be very suitable for quantitative analysis. However, they require calibration curves for each compound as each one gives a different detector response. For qualitative analysis, chromatographic methods (HPLC with diode array detector, MS, and/or NMR, GC–MS) are the most powerful as they offer both retention behavior and physical characteristics as a tool for identification. MS allows the determination of the molecular weight, and in case of high resolution also of the elemental composition, but this is not always sufficient to determine the structure. Tandem MS might be of help to identify, in such a case, the compound through its fragmentation pattern, but in case of novel compounds further spectral data are required. The performance for NMR can be improved by using two-dimensional (2D) NMR methods, which even may enable structure elucidation of novel compounds in a mixture.
(Verpoorte et al., 2007). In any case, the metabolite identification in non-targeted approaches in the absence of reference compounds remains difficult (Witzel et al., 2015), but there are different platforms and spectral databases online and it is expected that these difficulties could be solved in the future, as the number of identified compounds are increasing.

*Brassica* crops are attacked by an important number of pathogens which compromise the development of the plants of this genus. Some works have evaluated the effect of glucosinolates (secondary metabolites specific of brassicas) on different plant diseases (Ma et al., 2015; Singh et al., 2015; Sotelo et al., 2015; Velasco et al., 2013) but, despite of the importance of these crops and the effect of this pathogens, little is known about the interactions of different diseases on the metabolome of *Brassica* spp. Dr. Verpoorte’s lab (Leiden University) is the most active one working in *Brassica* metabolomics, using mainly NMR. So, till now, the only metabolome studies in *Brassica* crops, affected by diseases, were from or in collaboration with this lab. In 2008, Jahangir et al. (2008) evaluated the metabolomics response of *B. rapa* to different food borne bacteria such as *Bacillus subtilis, Staphylococcus aureus, Escherichia coli, Salmonella typhimurium and Shigella flexneri*. They used $^1$H NMR and two-dimensional NMR spectra, coupled with principal component analysis (PCA) and partial least square-discriminant analysis (PLS-DA). The metabolic changes varied according to bacterial species. Gram-positive and Gram-negative bacteria had a different effect on the *Brassica* metabolome. While threonine and GABA were found to be the discriminating metabolites in Gram-positive bacteria treated plants, Gram-negative bacteria exhibited a significant increase in sinapoyl-malate, caffeoyl-malate and histidine. In general, amino acids, alcohols, carbohydrates and phenols were the discriminating metabolites. Using the same species, Simoh et al. (2009) evaluated the
metabolomics changes that occur after infection with a disarmed and tumor-inducing strain of *Agrobacterium tumefaciens*. As in the previous report, the selected technique was $^1\text{H}$ NMR coupled with PLS. These authors observed differences between the two kinds of strains, affecting mostly to flavonoid, phenylpropanoid, sugar and free amino/organic acid content. As an interesting point, the infection with two tumor-inducing strains of *A. tumefaciens* provoke the suppression of some flavonoids and phenylpropanoids, compounds that are usually induced in the plants in response to biotic an abiotic stresses. In the same year, Abdel-Farid et al. (2009) evaluated the metabolite induction in *B. rapa* under the infection of several fungal pathogens (*L. maculans*, *Aspergillus niger*, and *F. oxysporum*). The results obtained from the $^1\text{H}$ NMR with PLS-DA, showed again that phenylpropanoids and phenols are associated with fungal infection. Besides, in this case, they found glucosinolates related to the infection response. Nevertheless, the metabolomic response was different for each fungal species, *F. oxysporum* infected plants accumulated more phenylpropanoids (sinapoyl-, feruloyl, and 5-hydroxyferuloyl malate), flavonoids (kaempferol and quercetin) and fumaric acid than plants infected with the other two species (Abdel-Farid et al., 2009).

In 2014, we started in our lab (Group of Genetics, Breeding and Biochemistry of Brassicas, MBG-CSIC) an experiment to evaluate the effect of *X. campestris* pv. *campestris* on the metabolome of *B. oleracea* at several times (1, 2, 3, 7, 14 days), using an LC-QTOF instrument. The preliminary results using the Welch test showed that several compounds are significantly affected by the infection and the time, both in negative and positive modes, from 94 compounds at day 1 to 300 at day 2, and decreasing during next days. Some of these compounds were present in the five times.

As far as we know, no other labs have investigated the metabolome of *Brassica* crops under the infection of different pathogens. In the model plants, *A. thaliana* and
Thellungiella, several authors have reported the metabolite profile with different diseases (Botanga et al., 2012; Floerl et al., 2012; Huang et al., 2009; Pedras and Zheng, 2010). In this way, Botanga et al. (2012) inoculated Col-0 *A. thaliana* with *A. brassicicola*. They showed that almost half of the detected compounds were affected by the inoculation. Some of these compounds, e.g. ascorbate, affect disease severity when applied to *Alternaria* in following experiments. *Verticillium longisporium* is one of the most devastating diseases in oilseed crops of the *Brassicaceae* and several authors have evaluated the changes in the metabolome under the infection of this fungus in *A. thaliana* (Floerl et al., 2012; König et al., 2014), by LC-QTOF or NMR. Several identified compounds (phytohormones, oxylipins, aminoacides, synapates, and lignans) and hundreds of potential biomarkers have been shown to be affected by infection of *V. longisporium*. The use of transformed *A. thaliana* to reduce the content of anti-nutritive compounds has led to discover of other metabolites implicated in disease response. In this way, Huang et al. (2009) transformed *A. thaliana* to reduce the content in sinapine. This transformation altered the metabolome resulting in changes of several sinapoyl derivatives, quercetin, salicylic acid and indolyl glucosinolates. The transformed plants were showed to be more susceptible to fungal infection indicating that these metabolites may have a role in the defense of *Brassicaceae* plants.

As showed above, few researches have been done to evaluate the role of the metabolome in the defense of *Brassica* diseases. From the data analyzed, it is clear that compounds from the phenylpropanoid route have an implication in the defense to several diseases but changes could occur in hundreds of metabolites and the response is plant and pathogen specific. Metabolomics is the newest ‘omic’ discipline and it has a broad field to develop. One of the key points will be the creation and improvement of
compound databases with MS and MS/MS information to identify relevant metabolic markers.

Conclusion

*Brassica* crops are susceptible to infections produced by fungus, bacteria or virus. These pathogens affect seriously the productivity and quality of these crops which result in substantial economic losses every year worldwide. For this reason, during the last decades it has been paid a great research emphasis on the plant host responses and the mechanisms underlying the resistance. Nevertheless, plant-pathogen interactions are highly complex since multiple pathogen factors and plant-signaling events take place, which ultimately define the susceptibility or resistance of the plant exposed to the pathogen. The recent advances in ‘omic’ approaches including gene expression analysis as well as protein and metabolite quantification enable genome-scale capturing of complex biological processes at the molecular level in plant diseases. This opens up new possibilities for understanding the molecular complexity of plant-pathogen systems and thus gain a better understanding of the molecular mechanisms implicated in basal and specific plant defense responses against a particular infection. To date, different ‘omic’ tools have been employed to understand how brassicas respond to biotic stress conditions. In this regard, genomics and transcriptomics have progressed as expected along with the recent availability of high-throughput sequence data generated from *Brassica* spp. but the other major ‘omic’ branches like proteomics and metabolomics are still lagging behind. In addition, the integration of ‘omic’-scale information to address complex genetics and physiological questions is still a challenge. Despite these limitations, the detailed global comparison of *Brassica* vegetables responses under
pathogen attack using ‘omics’ has allowed the identification of metabolic pathways, novel genes and proteins whose biological role warrants in-depth biochemical and cellular elucidation of resistance mechanism. Finally, a molecular-level understanding of biotic stress responses may identify promising novel targets for the development of *Brassica* cultivars with improved disease resistance.

**References**


homoeologous to the resistance gene cluster on *Arabidopsis* chromosome 4. Breed

Sci 62, 170-177.


