HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2016 May 01.

Published in final edited form as:

Nat Chem Biol. 2015 November ; 11(11): 837–839. doi:10.1038/nchembio.1914.

Two cytochromes P450 catalyze S-heterocyclizations in cabbage phytoalexin biosynthesis

Andrew P Klein and **Elizabeth S Sattely***

Department of Chemical Engineering, Stanford University, Stanford, California, USA

Abstract

Phytoalexins are abundant in edible crucifers and have important biological activities, yet no dedicated gene for their biosynthesis is known. Here, we report two new cytochromes P450 from *Brassica rapa* (Chinese cabbage) that catalyze unprecedented *S*-heterocyclizations in cyclobrassinin and spirobrassinin biosynthesis. Our results reveal the first genetic and biochemical insights into the biosynthesis of a prominent pair of dietary metabolites, and have implications for pathway discovery across >20 recently sequenced crucifers.

Cultivated and wild crucifers (Brassicaceae family) — e.g. broccoli (*Brassica oleracea*) and salt cress (*Eutrema salsugineum*, a wasabi relative) — produce an abundance of L-tryptophan-derived phytoalexins that function in the defense response against pathogens (Supplementary Results, Supplementary Fig. 1).[1,2] These compounds arise from the indole glucosinolate pathway[3] (Fig. 1) and have been linked to the health benefits of these vegetables; e.g. Chinese cabbage (*Brassica rapa*) produces brassinin, a suppressor of the PI3K/AKT/mTOR signaling pathway,[4,5,6] and cyclobrassinin, an antineoplastic[7]. Despite their unique chemical scaffolds, biological activities, and abundance in widely consumed plants (up to 5 mg in a 100 g fresh weight serving),[8] no dedicated gene for the biosynthesis of phytoalexins from edible crucifers is known.

The identification of phytoalexin biosynthetic genes would open the door to targeted breeding or metabolic engineering of these pathways for plant and human health. For example, the full elucidation of genes involved in glucosinolate biosynthesis in *Arabidopsis thaliana* enabled the pathway to be genetically engineered into tobacco[9] and yeast[10]. The link between glucosinolates and nutrition has motivated the commercialization of Beneforté broccoli[11] – a line bred for enhanced levels of glucoraphanin. In the context of plant health, glucosinolates[12,13] and camalexin, a phytoalexin made through a distinct biosynthetic pathway in Arabidopsis,[2] has exposed the role of small molecules in plant innate immunity.

*sattely@stanford.edu.

Accession codes. DDBJ/EMBL/GenBank: Nucleotide sequence data for *CYP71CR* genes in Brassicaceae species are available under the accession numbers TPA: BK009360–BK009375.. GEO: The transcriptome data have been deposited under accession number GSE69785.

Author contributions. A.P.K. and E.S.S. designed the study. A.P.K. performed all experiments. A.P.K. and E.S.S. analyzed the data and wrote the manuscript.

Competing Financial Interests Statement. The authors declare no competing financial interests.

The availability of plant genome sequences offers an unprecedented opportunity to accelerate the discovery of plant biosynthetic pathways; however, there is an urgent need to translate functional genomics tools from model organisms to crops and medicinal plants.[14]

The biosynthetic genes that comprise a pathway in plants can be dispersed across multiple chromosomes, rendering the discovery process slower than it is in microbial systems. Evidence of correlated expression can be a powerful way to associate genes into a pathway, but this approach is contingent on prior knowledge of one or more "bait" genes.[15] We reasoned that we could generate a list of candidate genes for phytoalexin biosynthesis in crucifers by merging condition-specific transcript sequencing with a biosynthetic hypothesis.

Indolic phytoalexins are produced by the recently sequenced Chiifu cultivar of *B. rapa*[16]; however, to increase experimental throughput, we surveyed in parallel the metabolism of a Rapid Cycling *B. rapa* (RCBr).[17] Since RCBr and Chiifu differ significantly in morphology and development, we first set out to determine whether phytoalexin biosynthesis is conserved across these two varieties of *B. rapa*. HPLC-MS-based metabolite profiling of Chiifu leaves revealed that three phytoalexins – brassinin, cyclobrassinin, and spirobrassinin (Fig. 1) – are produced when leaves are infiltrated with the pathogen *Pseudomonas syringae* pv. *maculicola* (*Psm*). (Supplementary Fig. 2). We observed similar accumulation of cyclobrassinin and spirobrassinin in response to *Psm* elicitation in RCBr leaves, while brassinin accumulates to a lesser extent (Fig. 2a). Furthermore, the conserved oligopeptide epitope of bacterial flagellin (flg22) is sufficient to induce the phytoalexin pathway in RCBr (Fig. 2a). Since RCBr retains the ability to make indolic phytoalexins, we chose this easy-to-cultivate variety for all subsequent experiments.

Previous isotope feeding studies suggested that brassinin is a precursor to cyclobrassinin and spirobrassinin.[18] We suspected that a cytochrome P450 (CYP) could catalyze oxidation and heterocyclization of brassinin to yield cyclobrassinin, since members of the CYP superfamily catalyze oxidations resulting in carbon–sulfur bond formation in glucosinolate biosynthesis[19] and camalexin biosynthesis[20]. To test for the presence of a CYP that could promote the formation of cyclobrassinin, we harvested RCBr leaves treated with *Psm* and isolated the microsomal protein fraction. We incubated leaf microsomes with brassinin and observed NADPH-dependent production of cyclobrassinin (Supplementary Fig. 3). These results are consistent with our hypothesis that a CYP is responsible for converting brassinin to cyclobrassinin in *B. rapa*.

To identify candidate genes responsible for the biosynthesis of cyclobrassinin, we used RNA sequencing (RNA-seq) to quantify transcripts extracted from RCBr leaves elicited with either *Psm* or flg22 in comparison to mock-treated leaves (Supplementary Table 1). Of the 353 CYP genes annotated in the *B. rapa* genome, we identified 15 genes in RCBr that were upregulated greater than four-fold following both flg22 and *Psm* treatments (Supplementary Table 2). Since indole glucosinolate biosynthetic genes are induced at least four-fold (Supplementary Table 3), we empirically chose this threshold for candidate gene selection (a two-fold cutoff would have yielded only five additional CYP candidates). Bioinformatic analysis of these upregulated CYP genes revealed that six of them (#1–#6) are homologous

to Arabidopsis genes that have a known function likely conserved in Brassica, seven genes (#7–#13) are homologous to Arabidopsis genes with unknown function, and two genes (*Bra009149* and *Bra005870*) have no clear homolog in Arabidopsis (Supplementary Table 4). Because the CYP superfamily has many examples of similar sequences that catalyze different reactions, we included in the list of candidates upregulated Brassica CYPs with an Arabidopsis homolog. We prioritized the candidates by first testing CYPs with no homolog in Arabidopsis (2 genes) or that were most similar to Arabidopsis CYPs of unknown function (7 genes) to increase the likelihood of discovering new enzymatic activity. Further bioinformatic analysis indicated that *Bra009149* and *Bra005870* are paralogs in *B. rapa*, together with a third gene *Bra009148* that had undetectable expression in the tested conditions. Of all annotated CYP genes in *B. rapa*, only these three paralogs lack a BLASTp hit in Arabidopsis with >40% identity, which suggests that they could serve a function that is not present in Arabidopsis.

To determine whether any of the enzymes with unknown function could oxidize brassinin into cyclobrassinin, five candidate genes (#7–#9, Bra009149, and Bra005870) were cloned, and expressed individually in *Saccharomyces cerevisiae* strain WAT11 together with CYP reductase *ATR1*.[21] Microsomes containing recombinant protein #7–#9 with NADPH were incapable of metabolizing brassinin (Supplementary Fig. 5). However, incubation of brassinin with Bra009149-containing microsomes resulted in NADPH-dependent conversion of brassinin to cyclobrassinin (Fig. 2b and Supplementary Fig. 6a). This result is consistent with a role for Bra009149 (renamed CYP71CR2) as a dedicated enzyme in the biosynthesis of cruciferous phytoalexins.

Surprisingly, Bra005870 (renamed CYP71CR1) also had NADPH-dependent catalytic activity on brassinin and produced a distinct product with a mass and MS/MS fragmentation pattern consistent with the structure of spirobrassinol (Fig. 2b and Supplementary Fig. 6b and 8). Only a trace amount of spirobrassinin was detected at extended incubation times suggesting that another enzyme is required to catalyze dehydrogenation of spirobrassinol in RCBr (Supplementary Fig. 7). These results indicate that CYP71CR1 catalyzes a distinct *S*-heterocyclization of brassinin to produce a spiro compound that is likely to be the immediate precursor to spirobrassinin.

We determined the kinetic parameters of CYP71CR2-catalyzed cyclobrassinin formation by HPLC-MS. The apparent K_m for brassinin is 248 ± 23 nM and apparent V_{max} is 478 ± 15 pmol mg^{-1} min^{-1} (Supplementary Fig. 10). CYP71CR1 also has an apparent K_m < 1 μM for brassinin (Supplementary Fig. 11). It was observed that CYP71CR2 and CYP71CR1 have slightly overlapping product profiles: 4% of the total products formed in the assay with CYP71CR1 is cyclobrassinin while 6% of the total products of CYP71CR2 is spirobrassinol (based on an estimated spirobrassinol extinction coefficient).

With biochemical evidence linking CYP71CR1 and CYP71CR2 to the biosynthesis of spirobrassinin and cyclobrassinin, respectively, we performed a timecourse to determine whether the expression of these genes correlates to metabolite production after elicitation (Supplementary Figs. 12–13). Each leaf, treated with either flg22 or *Psm*, was split in half and profiled by qRT-PCR for transcript abundance and by HPLC-MS for metabolite

quantification. We observed broad agreement between transcript and metabolite levels across the timecourse. In particular, *CYP71CR1* and spirobrassinin are both induced late and specifically by *Psm* elicitation.

The high amino acid sequence identity (70.2%) and similarity (85.6%) of CYP71CR2 and CYP71CR1 prompted us to consider the possibility that these enzymes oxidize brassinin by a similar mechanism. The formation of cyclobrassinin is likely to proceed by one of two routes, beginning with either hydroxylation at the 1′ position or epoxidation at the C2, C3 position (Fig. 2c). To discriminate between these two mechanisms, we used [1′, 1′-$^{2}H_{2}$]brassinin as a substrate in CYP71CR2 assays and found that >98% the label was retained in the product ([1′,1′-$^{2}H_{2}$]cyclobrassinin, Supplementary Fig. 14). This result rules out hydroxylation at the 1′ position, instead favoring a mechanism where CYP71CR2 catalyzes an initial epoxidation and *S*-heterocyclization at C2 to open the epoxide, followed by dehydration to restore indole aromaticity of cyclobrassinin (Supplementary Fig. 15). CYP71CR1 likely proceeds through the same epoxy-brassinin intermediate and *S*-heterocyclization at the C3 position to afford spirobrassinol. A similar mechanism was proposed in a recent report of spiro-carbon formation catalyzed by an unrelated flavin-dependent monooxygenase in the biosynthesis of spirotryprostatins in *Aspergillus fumigatus*.[22] Although indole epoxidation catalyzed by CYP71CR2 and CYP71CR1 results in a common intermediate en route to cyclobrassinin and spirobrassinol, we cannot rule out additional indole oxidation mechanisms (e.g. direct hydroxylation at C2) that could also lead to cyclobrassinin.

CYP71CR2, CYP71CR1, and their homologs constitute a new cytochrome P450 subfamily assigned the name CYP71CR.[23] We postulated that the presence of a *CYP71CR* member in a genome might be diagnostic of a species that produces phytoalexins derived from brassinin. As a first step to test this genotype–to–chemotype hypothesis, we searched the eighteen sequenced Brassicaceae genomes for *CYP71CR* genes, and compared the results with the reported phytoalexins isolated from cruciferous plants (Fig. 3). Consistent with our hypothesis, all plants known to produce brassinin-related phytoalexins contain at least one copy of the *CYP71CR* subfamily in their genome.[1,24] This association, for example in field pennycress (*Thlaspi arvense*) and salt cress (*Eutrema salsugineum*), suggests that these *CYP71CR* orthologs are involved in oxidative tailoring of brassinin scaffolds in these diverse species. A second inference from Figure 3 is that novel brassinin-derived phytoalexins may be produced in taxa that have not previously been chemically profiled, such as in London rocket (*Sisymbrium irio*), or an early diverging lineage of the crucifer family (*Aethionema arabicum*). In addition, phenylalanine-derived analogs of brassinin and cyclobrassinin were recently reported as phytoalexins in watercress[25] (*Nasturtium officinale*; same tribe as genome-sequenced *Leavenworthia alabamica*). We propose that these *CYP71CR* subfamily members identified in our analysis could be a key genetic determinant for the diversification of cabbage phytoalexins, and part of the arms race between plants and their fungal pathogens.

In summary, CYP71CR2 and CYP71CR1 are the first dedicated enzymes for phytoalexin biosynthesis in edible crucifers; these molecules likely have a role in both plant defense and human health through dietary exposure. CYP71CR1 and CYP71CR2 are the first

characterized members of a new P450 subfamily (CYP71CR) and catalyze differential processing of brassinin through unprecedented *S*-heterocyclizations. Phylogenetic distribution of *CYP71CR* genes among Brassicaceae species suggests that this CYP subfamily serves an important biological function, perhaps the construction of molecules that inhibit fungal pathogens. Efforts are underway to use coexpression and phylogenomic analysis to identify additional genes in the biosynthesis of brassinin and other cruciferous phytoalexins.

Online Methods

Materials and general methods

Synthetic standards were prepared according to literature procedures: brassinin was synthesized from 3-(aminomethyl)indole,[26] [1′,1′-$^2H_2$]brassinin was synthesized from 3-cyanoindole,[27] and both cyclobrassinin and spirobrassinin were synthesized from brassinin.[7,28] All synthesized compounds displayed NMR and high-resolution mass spectrometry (HRMS) spectral data consistent with those reported in the literature. ^{1}H NMR spectra were acquired on a Varian 400 MHz spectrometer. HRMS data were obtained using an Agilent 6520 Q-TOF, as described in the “Metabolite extraction and LC-MS analysis” section below. All other chemicals were obtained from Sigma-Aldrich or Thermo Fisher Scientific, unless otherwise stated. Ultrapure water was generated by a Milli-Q system (EMD Millipore).

B. rapa cv. Chiifu growth conditions and phytoalexin elicitation

Seeds of *Brassica rapa* ssp. *pekinensis* line Chiifu-401-42 were a generous gift from Dr. Beom-Seok Park at the National Academy of Agricultural Science, Korea. Each seed was germinated in a cylindrical pot (16 cm dia., 17 cm height) with Miracle-Gro Potting Mix (Scotts Company), under a 16-h photoperiod at room temperature (21–24°C). After 7 weeks, the plants were transferred to 4°C under a 16-h photoperiod for vernalization. Plants were elicited for phytoalexin production 12 weeks after germination, once 12–14 true leaves had emerged. For elicitation, two young and almost fully expanded leaves per plant (typically true leaves #8 and 9) were detached at the petiole using a razor blade to facilitate handling. A needle was used to lightly scratch the abaxial side of the leaf, and then a needleless 1-mL syringe was used to infiltrate the appropriate elicitor solution (described below). Each infiltrated leaf was placed in a covered petri dish (150 mm dia.) containing 75 mL sterile water to maintain high humidity, and incubated at 22°C under a 16-h photoperiod. At 48 hours post-infiltration, ≈2 cm^2 tissue samples were cut from each elicited leaf, placed in a pre-weighed 2-mL Safe-Lock Tube (Eppendorf), flash-frozen in liquid nitrogen, and stored at −80°C until metabolite extraction.

RCBr growth conditions and phytoalexin elicitation

Seeds of Rapid Cycling *Brassica rapa* (stock #1–33, rapid cycling ideotype) were kindly supplied by the Rapid Cycling Brassica Collection at the University of Wisconsin-Madison. Each seed was germinated in an 8-cm square pot with Pro-Mix HP (Premier Tech), in a 16-h photoperiod growth chamber (Percival Scientific) maintained at 22°C and 50% humidity. Plants were elicited for phytoalexin production at 15 days old, once 4–5 true leaves had

emerged. For elicitation, two young and almost fully expanded leaves per plant (typically true leaves #2 and 3) were infiltrated—using a needleless 1-mL syringe pressed against the abaxial side—with the appropriate elicitor solution (described below). Treated plants were returned to the growth chamber. At the indicated number of hours post-infiltration (hpi), each elicited leaf was detached at the petiole, placed in a pre-weighed 2-mL Safe-Lock Tube, flash-frozen in liquid nitrogen, and stored at −80°C until metabolite or transcript extraction.

One RCBr leaf (≈150 mg fresh weight) was treated as one biological replicate. Leaves designated for RNA-seq or qRT-PCR analysis (see below) were sectioned down the midrib, and each half was placed in a separate pre-weighed 2-mL Safe-Lock Tube prior to flash-freezing. For these paired samples, one half of the leaf was used for transcript analysis (RNA-seq or qRT-PCR), and the half was analyzed by LC-MS (see below). Biological replicates were randomized to control for (1) potential genetic variation in this self-incompatible population of RCBr, and (2) any distal signaling within the plant, i.e. from one infiltrated leaf to another. For a scheme of a typical experimental design, see Supplementary Figure 4.

Phytoalexin elicitors

Pseudomonas syringae pv. *maculicola* ES4326 (*Psm*) from a −80°C glycerol stock was streaked onto LB agar plates (with 100 μg/mL streptomycin) and grown at 30°C for 2 days (or for 1 day if streaking from a fresh liquid culture). The cells were suspended in 10 mM $MgCl_2$, pelleted at 10,000×g rcf, and resuspended in 10 mM $MgCl_2$ to a final OD_{600} of 0.01. For flg22 treatment, a sterile 10 μM solution of synthetic flg22 oligopeptide (N-term-QRLSTGSRINSAKDDAAGLQIA-C-term, Stanford PAN Facility) was prepared in 10 mM $MgCl_2$. For mock treatment, a sterile solution of 10 mM $MgCl_2$ was used. Leaf infiltrations were performed 3–4 h after the start of the light cycle, to control for effects of the circadian rhythm.

Metabolite extraction and LC-MS analysis

For metabolite extraction, frozen samples were lyophilized to dryness. Next, 100 μL of an 80:20 MeOH/H_2O (v/v) solution was added per milligram of dry tissue. The samples were subsequently homogenized on a ball mill (Retsch MM400) using 5 mm diameter steel beads, shaking at 25 Hz for 2 min. The tissue extracts were heat-treated at 65°C for 10 min, spun at 17,000×g rcf, 4°C, for 5 min, passed through 0.45 μm PTFE filters, and analyzed (10 μL injection volume) by LC-MS.

Liquid chromatography coupled to mass spectrometry (LC-MS) was performed on an Agilent 1260 HPLC, using a Gemini NX-C18 column (Phenomenex, 5 μm, 2 × 100 mm). Water and acetonitrile—each supplemented with 0.1% formic acid—were used as the mobile phase components, with a flow rate of 0.4 mL/min. The following 29-min gradient was used (percentages indicate acetonitrile concentration): 5–75% over 18 min; 75–95% over 1 min; 95% for 4 min; 95–5% over 1 min; 5% for 5 min. Column eluent was monitored by an Agilent 1260 diode array detector followed by an Agilent 6520 Accurate-Mass Q-TOF mass spectrometer with an ESI source (parameters: mass range: 50–1200 m/z; drying gas:

350°C, 11 L/min; nebulizer: 35 psig; capillary: 3000 V; fragmentor: 150 V; skimmer: 65 V; octupole 1 RF Vpp: 750 V; 1000 ms per spectrum). Mass spectrometry data were collected in positive ion mode and analyzed using MassHunter Qualitative Analysis software (Agilent). Phytoalexin content was calculated from the integrated peak area of a selected ion chromatogram (extracted using a ±10 ppm window around the theoretical exact mass given in Supplementary Table 6) and comparison to a standard curve.

RCBr leaf microsomal protein isolation

RCBr leaves were elicited with *Psm* and harvested 24 hpi, as described in the "RCBr growth conditions and phytoalexin elicitation" section above. The microsomal protein fraction was isolated based on the protocol of Schuhegger *et al*.,[29] with some modifications. In brief, 500 mg (fresh weight) of frozen tissue was ground in a mortar and pestle together with ~10 mg of poly(vinylpolypyrrolidone), ~40 mg of acid-washed sand, and 10 mL of ice-cold isolation buffer (100 mM Tris, pH 7.5; 100 mM ascorbic acid; 20% sucrose; 20% glycerol; 5 mM DTT; 1 mM EDTA; 1 mM PMSF). All of the subsequent steps were performed at 4°C. The homogenized sample was diluted with isolation buffer to a total volume of 30 mL, centrifuged twice at 15,000×g rcf for 10 min, discarding the pellet and retaining the supernatant each time. Microsomal protein was pelleted by ultracentrifugation at 200,000×g rcf for 40 min using a Beckman Coulter Optima L-90K ultracentrifuge, type 70 Ti rotor, and QuickSeal tubes. The pellet was resuspended in 0.5 mL of storage buffer (50 mM KP_i, pH 7.5; 20% glycerol; 1 mM DTT), frozen in liquid nitrogen, and stored at −80°C. Total protein content of the microsomes was estimated by absorbance at 280 nm using a NanoDrop 1000 spectrophotometer.

Detection of enzyme activity in the microsomal protein fraction

To a solution of NADPH (5 mM), brassinin (0.2 mM; from a 10 mM stock solution in DMSO), glutathione (1 mM), and sodium phosphate buffer (50 mM, pH 7.5) was added RCBr leaf microsomes (40 mg/mL) to initiate the reaction (total reaction volume of 200 μL). The resulting mixture was incubated at ambient temperature. At time points (12 min, 1 h, 4 h), the reaction was mixed by briefly vortexing, and a 50-μL aliquot was transferred to a tube containing 50 μL of MeOH to terminate enzyme activity. Samples were diluted with 100 μL of H_2O, passed through a 0.45-μm PFTE filter, and analyzed (20 μL injection volume) by LC-MS as described above. As negative controls, NADPH was omitted from the reaction mixture. Microsomal protein was isolated on two separate occasions, with no significant variation in results.

Next-generation RNA sequencing (RNA-seq)

Total RNA was isolated from RCBr leaf tissue using TRIzol Reagent (Life Technologies) according to the manufacturer's instructions. Inspection of total RNA using a BioAnalyzer 2100 (Agilent Technologies) indicated that the samples were of adequate RNA integrity (RIN ≥ 6.5). Library preparation was performed by the Stanford Functional Genomics Facility, using the TruSeq RNA Sample Prep Kit v2 (Illumina). The libraries were sequenced on the Illumina MiSeq platform (2×150 bp paired-end reads). Four conditions were included: flg22 (F), *Psm* (P), and mock-treated plants (10 mM $MgCl_2$; termed Mf and

Mp, respectively). Flg22 and Mf plants were grown side-by-side on one occasion, and *Psm* and Mp plants were grown side-by-side on a separate occasion. All samples for RNA-seq were harvested 9 hours post-infiltration (hpi), which coincides with the largest increase in cyclobrassinin accumulation (Supplementary Fig. 12b). Six biological replicates were first analyzed by LC-MS. The three biological replicates for each condition showing the greatest difference in phytoalexin content (for the flg22-vs.-Mf and *Psm*-vs.-Mp comparisons) were selected for RNA-seq analysis, using the paired sample collected at the time of harvest (see above). RNA-seq read quality and mapping statistics are displayed in Supplementary Table 1.

The RNA-seq data analysis pipeline consisted of four steps. First, the following quality control operations were performed using Trimmomatic[30] (v. 0.32): Illumina adapter sequences were removed, errors from random hexamer mispriming were excluded by cropping the first 10 bases from each read, low quality bases were removed, and only reads of 40 bases or longer were retained. Second, reads were aligned to the *B. rapa* reference genome[16] (IVFCAASv1, INSDC Assembly GCA_000309985.1) using TopHat[31] (v. 2.0.10) with the following modified parameters: minimum intron length was 40, maximum intron length was 5000, allowed up to 3 mismatches per read, and allowed read gaps of up to 3 bases. Third, transcripts were assembled using the Cufflinks[31] package (v. 2.1.1), enabling both multi-read correction and fragment bias correction, and quantified as Fragments Per Kilobase of exon per Million mapped fragments (FPKM). Fourth, the expression levels of all cytochrome P450 (CYP) genes—as annotated in Phytozome[32] (v. 9.1)—were extracted to a spreadsheet using a custom Python script. The transcriptome data have been deposited onto GEO (Accession code: GSE69785).

Bioinformatic analysis of CYP candidate genes

An *Arabidopsis thaliana–Brassica rapa* comparative genomic analysis was performed on the fifteen cytochrome P450 genes that were at least four-fold upregulated in both the flg22-vs.-Mf and *Psm*-vs.-Mp comparisons, as follows. Each amino acid sequence was used as a BLASTp query to search the database of Arabidopsis proteins (TAIR10). A BLASTp best hit with ≥80% pairwise identity was deemed homologous. The biochemical function, if known, of the Arabidopsis homologs was obtained from TAIR (http://arabidopsis.org). Two of the *B. rapa* proteins (Bra009149 and Bra005870) had no BLASTp hit with ≥40% pairwise identity, and were deemed to not possess a homolog in Arabidopsis. Bra009149 and Bra005870 were renamed CYP71CR2 and CYP71CR1, respectively, by the CYP nomenclature committee. A summary of the gene expression and Arabidopsis homology analysis is displayed in Supplementary Table 4.

Bioinformatic analysis of CYP71CR genes in Brassicaceae genomes

An expanded bioinformatic analysis was performed on the CYP71CR subfamily. DNA sequences of CYP71CR homologs were obtained using the tBLASTn program, with CYP71CR2 or CYP71CR1 as a protein query to search the Whole-Genome Shotgun contigs (WGS) database on NCBI (accessed on 10 October 2014; Supplementary Table 8). No CYP71CR sequence had been functionally characterized prior to this work, to the best of our knowledge. Coding sequences for fourteen additional CYP71CR members were found in

cruciferous plants (family: Brassicaceae), and these DNA and protein sequences have been deposited into GenBank (Accession codes: BK009360–BK009375). The protein sequences were aligned using Geneious (v. 7.1, using default parameters, see Supplementary Fig. 16). The phylogenetic tree was built using Geneious (v. 7.1) with a neighbor-joining method, no outgroup, and 1,000 bootstrap resamplings (see Supplementary Fig. 17). Phylogenetic relationships of crucifer species are based on BrassiBase.[33]

Cloning and heterologous production of CYP proteins in a yeast heterologous host

Libraries of cDNA were synthesized from the same RNA samples used for RNA-seq using SuperScript III First-Strand Synthesis kit (Life Technologies) according to the manufacturer's instructions. CYP-encoding sequences were amplified from the cDNA library by PCR using Phusion DNA Polymerase (Thermo Scientific) and oligonucleotide primers (Integrated DNA Technologies; sequences given in Supplementary Table 5). PCR products of the desired size were gel-purified (Zymo Research), inserted into the vector pYeDP60 (ref. 21) using the *Bam*HI and *Eco*RI restriction sites (enzymes from New England Biolabs), and transformed into *Escherichia coli* TOP10 (Life Technologies). Because RCBr is heterozygous at some genomic loci, eight individual clones of each *CYP* were purified (QIAprep kit from Qiagen), sequenced (Elim BioPharm), and compared. If two alleles for a particular *CYP* were present, then the allele with the higher % amino acid identity to the *B. rapa* Chiifu reference genome was chosen. The procedure for heterologous expression of CYPs using *Saccharomyces cerevisiae* strain WAT11 (kindly supplied by Dr. Franck Pinot)[21] was identical to that described in Klein *et al*.[20]

In vitro assay of recombinant CYP enzymes

To a solution of NADPH (1 mM), brassinin (10 μM; from a 0.5 mM stock solution in DMSO), and sodium phosphate buffer (50 mM, pH 7.5) was added yeast microsomes harboring a CYP (1.5 mg/mL) to initiate the reaction (total reaction volume of 150 μL). The resulting mixture was incubated at 30°C for 15 min. The reaction was stopped by the addition of 60 μL of MeOH, passed through a 0.45-μm PFTE filter, and analyzed (20 μL injection volume) by LC-MS. The microsome fraction isolated from yeast harboring the empty pYeDP60 vector was used as a negative control. For CYPs that did not show catalytic activity on brassinin at 10 μM, additional assays using brassinin at 200 μM were performed, but no significant substrate consumption or product formation were observed.

Steady-state kinetic analysis of CYP enzymes

For steady-state kinetic analysis of CYP71CR2 (Bra009149), a solution of NADPH (1 mM), sodium phosphate buffer (50 mM, pH 7.5), and yeast microsomes harboring CYP71CR2 (0.2 mg/mL) was prepared at 30°C (total volume of 1470 μL). To initiate the reaction, 30 μL of a brassinin stock solution (in DMSO) was added, the components were mixed by briefly vortexing, and then returned to incubate at 30°C. At each time point, the reaction was briefly vortexed, and a 150-μL aliquot was transferred to a cold tube (−20°C) containing 50 μL of MeOH (with 1 μM camalexin as an internal standard) to terminate enzyme activity. Samples were passed through a 0.45-μm PFTE filter and analyzed (20 μL injection volume) by LC-MS. Construction of standard curves for brassinin and cyclobrassinin using authentic

standards (Supplementary Fig. 9) allowed us to calculate the concentration of substrate and product in each assay sample. The cyclobrassinin concentrations in the eight time points (1, 2, 4, 6, 8, 10, 12, 15 min) were fit to a one-phase decay curve using GraphPad Prism software (v. 6.0d). The initial rate of reaction was determined from the first derivative of the exponential function. The following eleven brassinin final concentrations were each tested on three separate occasions: 50, 75, 100, 150, 200, 300, 400, 600, 800, 1200, 1600 nM. The kinetic constants for brassinin, apparent K_m and V_{max}, were inferred using a nonlinear regression in GraphPad Prism (Supplementary Fig. 10).

The initial velocity of CYP71CR1 (Bra005870) was determined by standard end-point enzyme activity assays similar to those described above for CYP71CR2, except microsomes harboring CYP71CR1 (0.3 mg/mL) were used instead. Attempts to isolate or synthesize spirobrassinol for NMR spectroscopic analysis were unsuccessful, as the compound degraded during purification. To overcome the instability of spirobrassinol, the hemiaminal was trapped by the addition of excess cysteine (40 mM; from a 0.8 M stock solution) after the methanol quench. Following a 10 minute incubation at room temperature, the peak corresponding to spirobrassinol had disappeared, and a new peak corresponding to the spirobrassinol–cysteine adduct arose, which was stable over the duration (>12 hours) of a typical kinetics experiment. The formation of the spirobrassinol–cysteine adduct was monitored by absorption at 283 nm (34 nm bandwidth). The relative peak areas in the eight time points were used to determine initial rate of reaction, as described above. The following four brassinin concentrations were tested: 2, 10, 25, 50 μM. No significant difference in the initial velocity was observed, so an apparent $K_m < 1$ μM was inferred (Supplementary Fig. 11).

The following variations to the “LC-MS analysis” section were used when analyzing samples for steady-state kinetics: column: Poroshell 120 EC-C18 (Agilent, 2.7 μm, 3 × 50 mm); flow rate: 0.5 mL/min; drying gas: 12 L/min; mass range: 80–1000 m/z; time segment: 667 ms per spectrum; and the following 13.5-min gradient was used: 17–77% over 8 min; 77–95% over 0.1 min; 95% for 2 min; 95–17% over 0.4 min; 17% for 3 min.

Gene transcript analysis by quantitative real-time PCR (qRT-PCR)

RCBr leaves were frozen in liquid nitrogen, ground using a pre-chilled mortar and pestle, and subjected RNA extraction using the Spectrum Plant Total RNA Kit (Sigma), including on-column DNaseI digestion. RNA (2.5 μg) was used to synthesize cDNA in 20-μL reactions using SuperScript III First-Strand Synthesis Kit (Life Technologies) with oligo(dT)^{20} primers. The cDNA served as template for qRT-PCR performed using the StepOnePlus Real Time PCR System (Applied Biosystems). Each reaction (total volume 15 μL) contained 200 nM of each primer, 4 ng/μL of cDNA (based on total RNA input), and 1X of SensiMix SYBR Hi-ROX reagent mix (Bioline).

The Gene-specific primer sequences are given in Supplementary Table 7. The amplification efficiency of each primer pair was between 93% and 98%, as determined by testing serial dilutions of plasmid DNA containing the cloned, Sanger-sequenced genes (with the exception of *CYP71CR3*, *ACTIN2*, and *GAPDH*, where the primer pairs were validated using cDNA serial dilutions). Cross-hybridization of *CYP71CR* primer pairs between one

another was negligible; C_T values were >18 cycles higher when using a different plasmid DNA template, compared to the matching plasmid DNA template. As reference genes, we tested *ACTIN2* (*Bra022356*) and *GAPDH* (*Bra016729*), which have each been used successfully in previous qRT-PCR studies of *B. rapa*.[34,35] The two reference genes yielded essentially identical normalized transcript levels for the target genes, and results are presented using *ACTIN2* as the reference gene. Fold-change was calculated using the $\Delta\Delta C_T$ method. Four biological replicates, each consisting of a paired sample (see above), were included. One samples was analyzed by LC-MS, and the other was analyzed by qRT-PCR with three technical replicates (Supplementary Figs. 12–13).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by the US National Institutes of Health grants R00 GM089985 and DP2 AT008321 (E.S.S.). A.P.K. was supported by the NDSEG Fellowship. We thank M. Jin (Korea Brassica Genome Project) and H.B. Lauffer (Wisconsin Fast Plants Program) for providing seeds. We thank X. Ji (Stanford Functional Genomics Facility) for RNA sequencing services. We thank L. Meyer-Teruel for help in initial metabolic profiling of *B. rapa*, J.-G. Kim and L. Sassoubre for qRT-PCR assistance, D. Nelson for CYP nomenclature assignment, and S. Galanie and members of the Sattely Group for helpful comments on the manuscript.

References for main text

1. Pedras MSC, Yaya EE, Glawischnig E. Nat. Prod. Rep. 2011; 28:1381–1405. [PubMed: 21681321]
2. Bednarek P. Chem Bio Chem. 2012; 13:1846–1859.
3. Pedras MSC, Yaya EE, Hossain S. Org. Biomol. Chem. 2010; 8:5150–5158. [PubMed: 20848032]
4. Banerjee T, et al. Oncogene. 2008; 27:2851–2857. [PubMed: 18026137]
5. Izutani Y, Yogosawa S, Sowa Y, Sakai T. Int. J. Oncol. 2012; 40:816–824. [PubMed: 22307336]
6. Kim S-M, et al. Phytother. Res. 2014; 28:423–431. [PubMed: 23686889]
7. Mehta RG, et al. Carcinogenesis. 1995; 16:399–404. [PubMed: 7859373]
8. Takasugi M, Monde K, Katsui N, Shirata A. Bull. Chem. Soc. Jpn. 1988; 61:285–289.
9. Pfalz M, et al. Plant Cell. 2011; 23:716–729. [PubMed: 21317374]
10. Mikkelsen MD, et al. Metab. Eng. 2012; 14:104–111. [PubMed: 22326477]
11. Traka MH, et al. New Phytol. 2013; 198:1085–1095. [PubMed: 23560984]
12. Clay NK, Adio AM, Denoux C, Jander G, Ausubel FM. Science. 2009; 323:95–101. [PubMed: 19095898]
13. Bednarek P, et al. Science. 2009; 323:101–106. [PubMed: 19095900]
14. Bradbury LM, Niehaus TD, Hanson AD. Curr. Opin. Biotechnol. 2013; 24:278–284. [PubMed: 22898705]
15. Saito K, Hirai MY, Yonekura-Sakakibara K. Trends Plant Sci. 2008; 13:36–43. [PubMed: 18160330]
16. Wang X, et al. Nat. Genet. 2011; 43:1035–1040. [PubMed: 21873998]
17. Williams PH, Hill CB. Science. 1986; 232:1385–1389. [PubMed: 17828914]
18. Monde K, Takasugi M, Ohnishi T. J. Am. Chem. Soc. 1994; 116:6650–6657.
19. Bak S, Feyereisen R. Plant Physiol. 2001; 127:108–118. [PubMed: 11553739]
20. Klein AP, Anarat-Cappillino G, Sattely ES. Angew. Chem. Int. Ed. 2013; 52:13625–13628.
21. Pompon D, Louerat B, Bronine A, Urban P. Methods Enzymol. 1996; 272:51–64. [PubMed: 8791762]
22. Tsunematsu Y, et al. Nat. Chem. Biol. 2013; 9:818–825. [PubMed: 24121553]

23. Nelson DR. Hum. Genomics. 2009; 4:59–65. [PubMed: 19951895]
24. Bednarek P, et al. New Phytol. 2011; 192:713–726. [PubMed: 21793828]
25. Pedras MSC, To QH. Phytochemistry. 2015; 113:57–63. [PubMed: 25152450]

Methods-only references

26. Pedras MSC, Jha M. Bioorg. Med. Chem. 2006; 14:4958–4979. [PubMed: 16616505]
27. Pedras MSC, Minic Z, Sarma-Mamillapalle VK. Bioorg. Med. Chem. 2011; 19:1390–1399. [PubMed: 21292494]
28. Pedras MSC, Suchy M, Ahiahonu PWK. Org. Biomol. Chem. 2006; 4:691–701. [PubMed: 16467943]
29. Schuhegger R, et al. Plant Physiol. 2006; 141:1248–1254. [PubMed: 16766671]
30. Bolger AM, Lohse M, Usadel B. Bioinformatics. 2014; 30:2114–2120. [PubMed: 24695404]
31. Trapnell C, et al. Nat. Protoc. 2012; 7:562–578. [PubMed: 22383036]
32. Goodstein DM, et al. Nucleic Acids Res. 2012; 40:D1178–D1186. [PubMed: 22110026]
33. Kiefer M, et al. Plant Cell Physiol. 2014; 55:e3. [PubMed: 24259684]
34. Wiesner M, Schreiner M, Zrenner R. BMC Plant Biol. 2014; 14:124. [PubMed: 24886080]
35. Qi J, et al. Plant Mol. Biol. Report. 2010; 28:597–604.

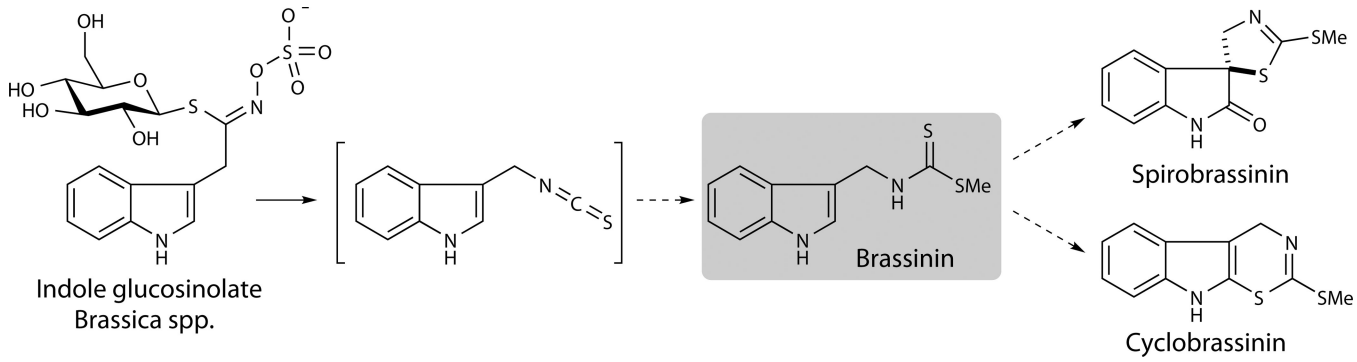

Figure 1. Phytoalexins in *Brassica rapa*

The biosynthetic relationship of indole glucosinolate, brassinin, cyclobrassinin, and spirobrassinin produced by cruciferous plants. The solid arrow indicates a reaction catalyzed by a known enzyme, and dotted arrows indicate putative enzymatic reactions supported by previous labelling studies.[1,18]

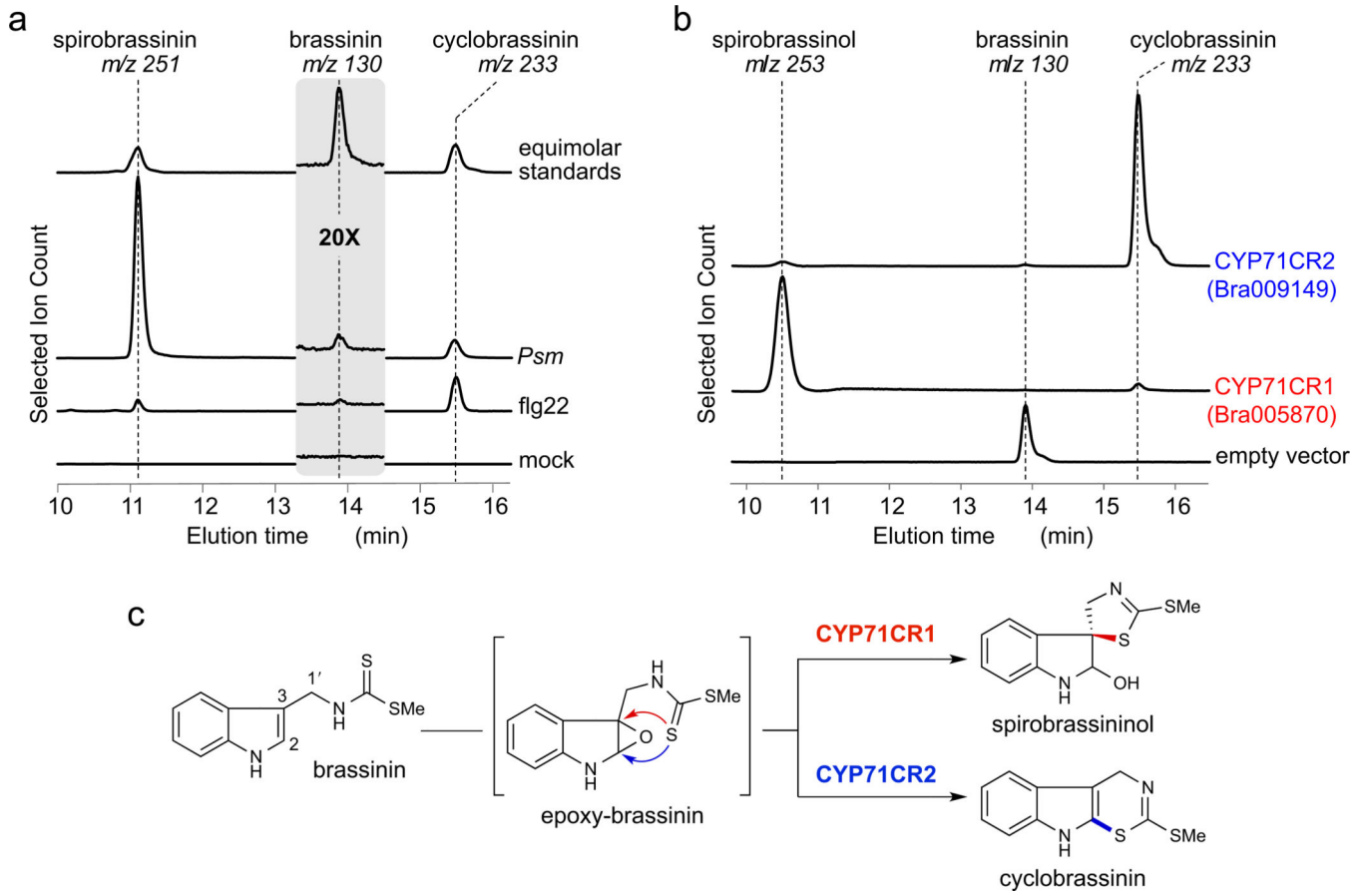

Figure 2. Phytoalexin production in *B. rapa* and *in vitro* analysis of cytochromes P450 in cruciferous phytoalexin biosynthesis

(**a**) HPLC-MS analysis of metabolites extracted from Rapid Cycling *B. rapa* (RCBr) leaves harvested at 24 hours post-infiltration. Traces show selected ion monitoring of *m/z* 130, 233, and 251 for brassinin, cyclobrassinin, and spirobrassinin, respectively. Standards (10 pmol of each) are displayed to show differences in ionization efficiency. The chromatogram is magnified 20-fold in the vertical direction around the elution time of brassinin to show detail. The chromatograms are representative of experiments was performed with biological replicates (n = 4) on three separate occasions. *Psm*, *Pseudomonas syringae* pv. *maculicola*; flg22, conserved oligopeptide epitope of bacterial flagellin; mock, sterile buffer. (**b**) HPLC-MS analysis of the reaction products generated by incubating brassinin with yeast-expressed CYP71CR2 (Bra009149) or CYP71CR1 (Bra005870) *in vitro*. Traces show selected ion monitoring of *m/z* 130, 233, and 253 for brassinin, cyclobrassinin, and spirobrassinol, respectively. The chromatograms are representative of reactions prepared on three separate occasions. (**c**) Reactions catalyzed by CYP71CR1 and CYP71CR2, with the proposed intermediate epoxy-brassinin. Spirobrassinol is likely the immediate precursor to spirobrassinin.

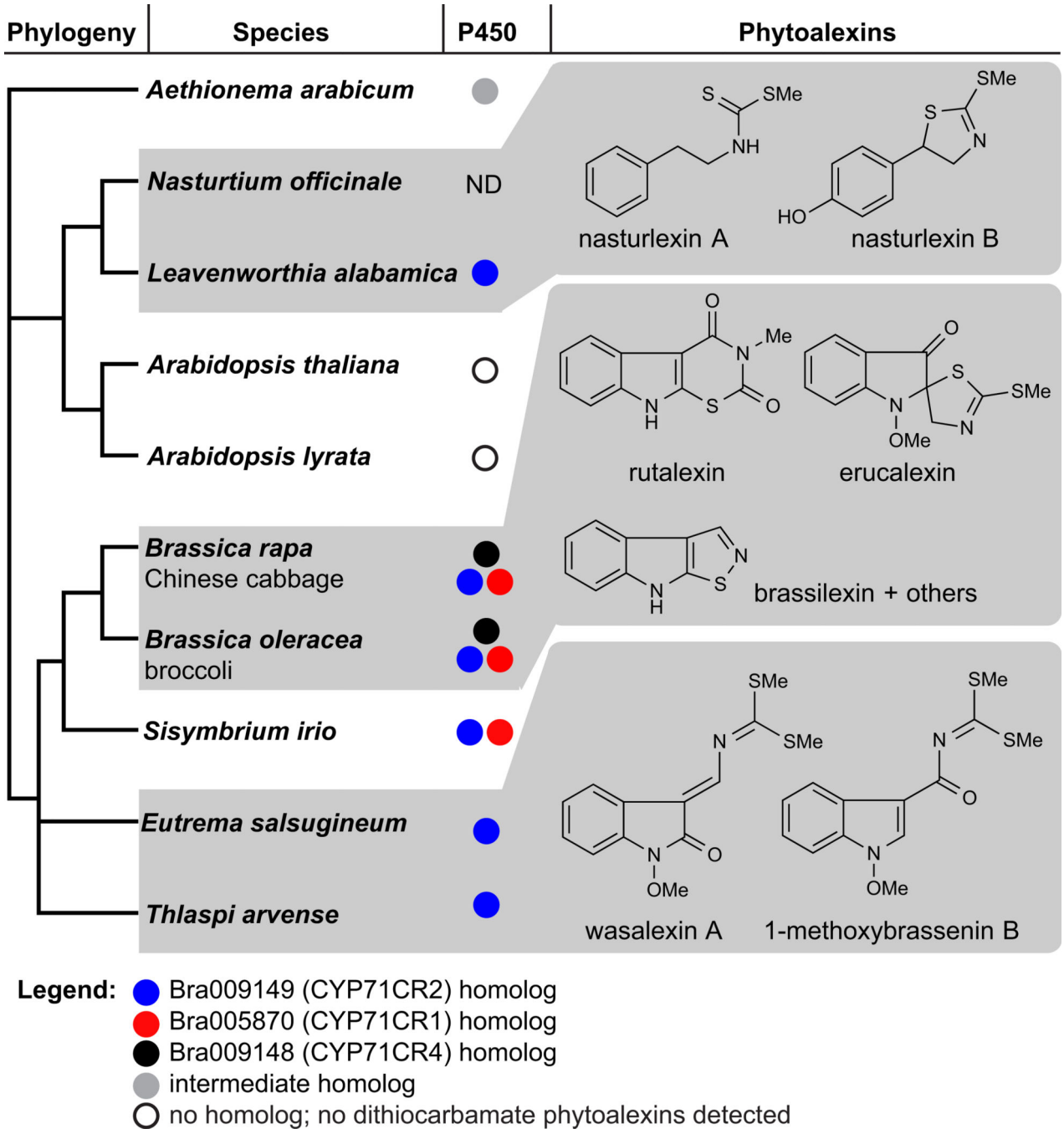

Figure 3. Occurrence of *CYP71CR* genes and phytoalexins across the Brassicaceae family

Phylogenetic relationships of crucifer species; branch lengths are not to scale. The species displayed here have an available genome assembly (Supplementary Table 8), with the exception of *Nasturtium officinale*. Based on the close phylogenetic relationship of *N. officinale* and *Leavenworthia alabamica*, it is plausible that *N. officinale* has a *CYP71CR* gene. A phylogenetic analysis of CYP71CR members (Supplementary Figs. 16–17) was used to infer homology for these genes. Shading indicates the phytoalexins that are

characteristic of the particular clade of species, based on the recent literature on cruciferous phytoalexins.[1,24,25]