

# SCIENTIFIC REPORTS



OPEN

## Bioinformatic and expression analysis of the *Brassica napus* L. cyclophilins

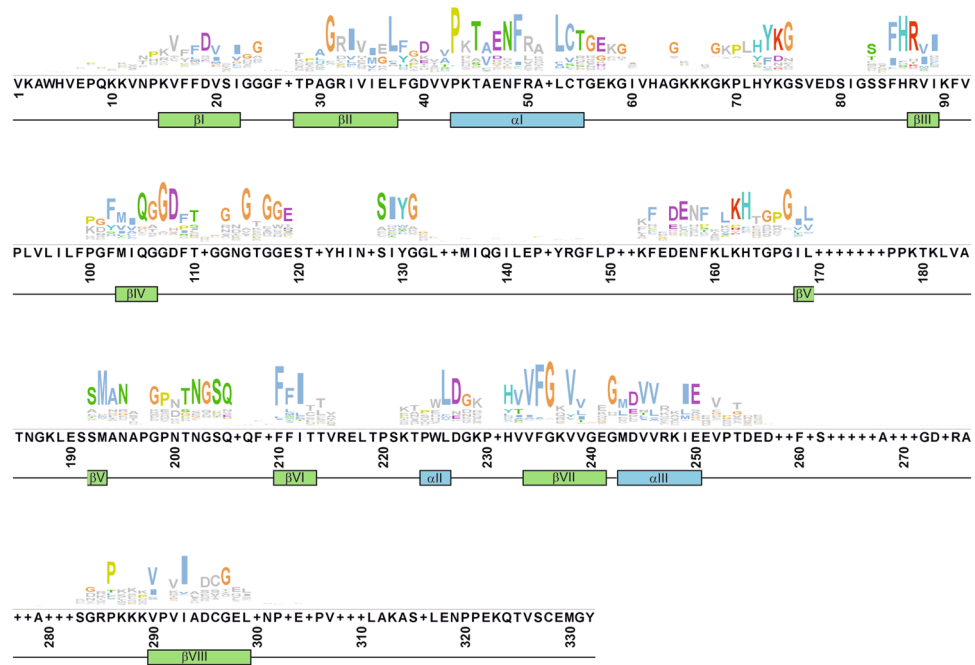
Patrizia Hanhart<sup>1</sup>, Melanie Thieß<sup>1</sup>, Khalid Amari<sup>2</sup>, Krzysztof Bajdzienko<sup>3</sup>, Patrick Givalisco<sup>3</sup>, Manfred Heinlein<sup>2</sup> & Julia Kehr<sup>1</sup>

Cyclophilins (CYPs) are a group of ubiquitous proteins characterized by their ability to bind to the immunosuppressive drug cyclosporin A. The CYP family occurs in a wide range of organisms and contains a conserved peptidyl-prolyl *cis/trans* isomerase domain. In addition to fulfilling a basic role in protein folding, CYPs may also play diverse important roles, e.g. in protein degradation, mRNA processing, development, and stress responses. We performed a genome-wide database survey and identified a total of 94 CYP genes encoding 91 distinct proteins. Sequence alignment analysis of the putative BnCYP cyclophilin-like domains revealed highly conserved motifs. By using RNA-Seq, we could verify the presence of 77 BnCYP genes under control conditions. To identify phloem-specific BnCYP proteins in a complementary approach, we used LC-MS/MS to determine protein abundances in leaf and phloem extracts. We detected 26 BnCYPs in total with 12 being unique to phloem sap. Our analysis provides the basis for future studies concentrating on the functional characterization of individual members of this gene family in a plant of dual importance: as a crop and a model system for polyploidization and long-distance signalling.

As one of the most important crops for nutritional oil, fodder, biodiesel, chemical and pharmaceutical products, *Brassica napus* (oilseed rape or canola) is widespread in agriculture, especially in the European Union, China, and Canada. Beside its essential agricultural significance, *B. napus* has also become a model plant for studying long-distance signalling<sup>1–5</sup> and evolutionary consequences of polyploidization<sup>6,7</sup>. *Brassica napus* is used as a model plant for studying long-distance signalling, because methods for the collection of sufficient quantities of pure xylem and phloem sap are well established<sup>2,4</sup>. Compared to crops like wheat, soybean or rice, the domestication of *B. napus* was more recent. It is assumed that chromosome doubling occurred after spontaneous hybridization between *Brassica rapa* (Asian cabbage or turnip rape,  $2n = 2 \times 10 = 20$ , genome AA) and *Brassica oleracea* (Mediterranean cabbage,  $2n = 2 \times 9 = 18$ , genome CC)<sup>8</sup>. The profitable cultivation of *B. napus* (genome AACC,  $2n = 38$ ) has high importance for the economy<sup>9</sup>.

A group of proteins involved in diverse fundamental cellular functions in many different organisms is called immunophilins, originally discovered as receptors for immunosuppressive drugs in mammals. The family of immunophilins consists of two major groups, FK506-binding proteins (FKBPs)<sup>10,11</sup> and cyclophilins (CYPs)<sup>12,13</sup>. Despite the lack of sequence and structure similarity, FKBPs and CYPs each possess a conserved domain responsible for their common peptidyl-prolyl *cis/trans* isomerase (PPIase) activity, catalyzing the rate-limiting rotation of X-proline peptide bonds from a *cis* to a *trans* conformation<sup>14</sup>. These domains are called the FKBP and cyclophilin-like domain (CLD), respectively. An additional group of proteins exhibiting PPIase activity, the parvulins, cannot be classified as immunophilins in a strict sense, since they do not bind to any known immunosuppressant molecule<sup>15</sup>. The drugs of CYPs and FKBPs are cyclosporin A (CsA) and FK506/rapamycin, respectively, that bind to the catalytic pocket of the PPIase domain<sup>16</sup>, thereby inhibiting its activity and forming a high-affinity binding site for the interaction with calcineurin<sup>17</sup>. However, these drugs do not occur naturally in cells, therefore the consequences of drug treatment have clinical but no physiological relevance<sup>18</sup>.

<sup>1</sup>Molecular Plant Genetics, Universität Hamburg, Biozentrum Klein Flottbek, Ohnhorststraße 18, 22609, Hamburg, Germany. <sup>2</sup>Université de Strasbourg, CNRS, IBMP UPR 2357, 12 rue du Général Zimmer, F-67000, Strasbourg, France. <sup>3</sup>Max-Planck-Institut für Molekulare Pflanzenphysiologie, Wissenschaftspark Potsdam-Golm, Am Mühlenberg 1, 14476, Potsdam, Germany. Correspondence and requests for materials should be addressed to J.K. (email: [julia.kehr@uni-hamburg.de](mailto:julia.kehr@uni-hamburg.de))



**Figure 1.** Conserved sequence of the *B. napus* cyclophilin-like domain. CYP protein sequences were cropped to the predicted sequence for the CLD and aligned with the secondary structure of hCypA. The conserved sequence of 91 predicted CLDs highlights structural features and functional residues. Gaps represent insertions in very few members of this protein family.

Human CYPs were first characterized by their ability to bind the drug cyclosporin A in 1984<sup>12</sup>. The first plant CYPs were described in 1990, where CYP cDNA sequences were identified in tomato (*Lycopersicon esculentum*), maize (*Zea mays*), and oilseed rape (*B. napus*)<sup>19</sup>. The ubiquitous CYPs show highly conserved structural features<sup>20</sup> and are involved in several fundamental cellular functions, including protein folding<sup>21–23</sup>, protein trafficking<sup>24</sup>, signalling<sup>25–29</sup>, pathogen response<sup>30</sup>, apoptosis<sup>31</sup>, RNA-binding, regulation of gene expression or transcription<sup>32–36</sup>, and plant stress responses<sup>37–40</sup>.

Interestingly, CYPs were found to be a prominent and abundant class of proteins in the phloem long-distance system of higher plants including castor bean and oilseed rape<sup>2, 41</sup>. Phloem sieve elements lose their ability for transcription and translation during their maturation into transport tubes<sup>42</sup>. However, phloem exudate contains a complex set of proteins, some of which have been implicated with long-distance signalling<sup>43</sup>. Such signalling proteins are imported from the adjacent companion cells through plasmodesmata. It was proposed that phloem CYPs might act as molecular chaperones in this process, potentially involved in refolding of non-cell-autonomous proteins after entry into the translocation stream<sup>2, 41</sup>.

The growing number of sequenced genomes allowed the identification of whole sets of CYPs in various organisms by sequence comparisons. 19 CYPs were detected in the human genome, whereas *Saccharomyces cerevisiae* possesses 8, *Schizosaccharomyces pombe* 9, *Caenorhabditis elegans* 17, *Drosophila melanogaster* 14<sup>44, 45</sup>, and the fungus *Leptosphaeria maculans* 12<sup>46</sup> CYPs. Compared to other organisms, plants have a higher number of CYPs with 29 encoded in *A. thaliana*<sup>47–49</sup>, and 27 in rice<sup>37</sup>. To date, soybean (*Glycine max*) is reported to have the largest set of CYPs with 62 members<sup>50</sup>.

Based on the recently sequenced *B. napus* genome<sup>51</sup>, the major aim of the present study was the identification and classification of CYP-like proteins in this economically important species. A total number of 94 genes belonging to the CYP gene family (resulting in 91 different CYP proteins) could be identified in the sequenced *B. napus* cultivar ‘Darmor-bzh’. By applying transcriptome analysis of the *B. napus* cultivar ‘Drakkar’ we could confirm the expression of 77 *Bn*CYP genes in leaves. To identify CYP proteins specifically occurring in phloem sap, we performed complementary protein analyses by LC-MS/MS on leaf and phloem extracts and found 26 different *Bn*CYP proteins in total, 12 only present in phloem sap.

## Results and Discussion

**Identification of cyclophilins in the *Brassica napus* genome.** Putative CYPs of *B. napus* containing full length or partial CYP-like domains were identified by BLASTp of *A. thaliana* CYPs<sup>47, 48</sup>. The 94 *B. napus* (cultivar ‘Darmor-bzh’) gene sequences determined by this approach resulted in 91 distinct proteins which were subjected to additional analysis like sequence alignments with the respective Arabidopsis homologs and the verification of CLDs. The analysis showed that *B. napus* contains the largest CYP gene family known to date with 94 genes followed by soybean (*Glycine max*) with 62 members<sup>50</sup>.

As proposed by Romano *et al.*<sup>47</sup> and He *et al.*<sup>48</sup> we used the nomenclature *Bn*CYP (*Brassica napus* cyclophilin) followed by the molecular weight and a consecutive number for genes encoding proteins with similar molecular

Name	Gene Name	GenBank Accession	kDa	aa	CLD	pI (theoretical)	Exons
BnCYP5	BnaA01g25420D	CDY37322.1	5.7	49	2–42	8.7	1
BnCYP7-1	BnaA09g37110D	CDY06069.1	7.2	66	1–66	10.7	3
BnCYP7-2	BnaA01g36790D	CDY60385.1	7.3	64	1–61	4.5	1
BnCYP7-3	BnaA01g36800D	CDY60386.1	7.3	65	1–64	4.9	1
BnCYP8	BnaCnng42430D	CDY63671.1	8.2	71	9–65	9.1	2
BnCYP10-1	BnaC02g14160D	CDX96235.1	10.1	92	6–55	4.7	2
BnCYP10-2	BnaC06g10280D	CDY30664.1	10.5	94	1–68	6.7	2
BnCYP12-1	BnaC02g10590D	CDY02830.1	12.5	109	52–96	9.8	5
BnCYP12-2	BnaAnng41240D	CDY72459.1	12.8	117	43–116	9.0	3
BnCYP13	BnaA02g07550D	CDY32693.1	13.1	116	8–64	8.3	4
BnCYP14-1	BnaC04g41450D	CDX97794.1	14.0	127	1–123	5.6	1
BnCYP14-2	BnaC04g41430D	CDX97796.1	14.2	127	1–124	5.8	1
BnCYP14-3	BnaC04g41440D	CDX97795.1	14.2	127	1–123	6.0	1
BnCYP14-4	BnaC08g28870D	CDX71987.1	14.3	131	1–109	6.1	1
BnCYP14-5	BnaC09g34640D	CDX80278.1	14.4	133	84–132	9.4	3
BnCYP14-6	BnaA01g25470D	CDY37327.1	14.9	138	2–137	4.6	3
BnCYP16	BnaA02g10200D	CDY28828.1	16.1	148	3–124	4.4	3
BnCYP17-1	BnaA01g25460D	CDY37326.1	17.3	161	68–134	7.7	3
BnCYP18-1	BnaC04g09170D/BnaA05g08140D	CDX75068.1/CDX84155.1	18.2	164	7–163	8.9	5
BnCYP18-2	BnaC01g03590D	CDX69094.1	18.2	172	1–171	8.9	1
BnCYP18-3	BnaA01g02340D	CDX75468.1	18.2	172	1–172	8.9	1
BnCYP18-4	BnaA08g16920D/BnaC03g60160D	CDX90292.1/CDY12100.1	18.3	172	2–172	8.3	1
BnCYP18-5	BnaA09g08780D	CDY47469.1	18.4	171	1–171	8.3	1
BnCYP18-6	BnaA06g37360D	CDY22414.1	18.4	172	2–171	8.3	1
BnCYP18-7	BnaC07g47630D	CDX72741.1	18.4	172	2–171	8.3	1
BnCYP18-8	BnaC09g09060D	CDY26779.1	18.4	171	1–171	8.3	1
BnCYP19-1	BnaA09g35540D	CDY27248.1	19.6	182	1–174	7.7	1
BnCYP19-2	BnaC08g26990D	CDX72175.1	19.7	183	1–174	7.0	1
BnCYP20	BnaA01g36700D	CDY70394.1	20.9	192	15–191	8.2	1
BnCYP21-1	BnaCnng08980D	CDY38037.1	21.0	193	22–191	9.2	5
BnCYP21-2	BnaC04g06640D	CDX91486.1	21.6	200	23–200	8.2	7
BnCYP21-3	BnaAnng15590D	CDY58941.1	21.7	201	28–199	8.9	6
BnCYP21-4	BnaC03g12390D	CDX71155.1	21.7	201	26–199	9.3	7
BnCYP21-5	BnaC04g54560D	CDY55458.1	21.7	201	28–199	8.9	6
BnCYP21-6	BnaAnng17350D	CDY61026.1	21.8	201	26–199	9.3	7
BnCYP21-7	BnaA05g06380D	CDX74892.1	21.8	204	27–204	8.2	7
BnCYP21-8	BnaC04g45890D	CDX93307.1	21.9	205	28–205	8.2	6
BnCYP21-9	BnaC09g34020D/BnaA10g29520D	CDX80340.1/CDY52123.1	22.0	204	31–202	8.9	7
BnCYP22-1	BnaCnng32070D	CDY57430.1	22.0	191	133–176	6.5	4
BnCYP22-2	BnaA04g22160D	CDY18371.1	22.3	207	26–207	8.1	7
BnCYP22-3	BnaC04g41460D	CDX97793.1	23.0	215	55–212	5.7	2
BnCYP24-1	BnaC08g26840D	CDX72190.1	24.0	224	53–222	6.5	8
BnCYP24-2	BnaA09g35470D	CDY27255.1	24.0	224	53–222	6.5	8
BnCYP24-3	BnaA10g12350D	CDY33029.1	24.3	222	86–140	9.2	4
BnCYP24-4	BnaA08g10930D	CDX76307.1	24.4	223	42–215	6.5	7
BnCYP24-5	BnaC01g03530D	CDX69100.1	24.6	224	44–215	7.1	7
BnCYP24-6	BnaA01g02260D	CDX75476.1	24.6	224	43–215	7.1	7
BnCYP24-7	BnaC03g65620D	CDY17501.1	24.6	225	44–217	6.2	7
BnCYP25-1	BnaA08g20010D	CDY46429.1	25.4	226	26–198	8.4	8
BnCYP25-2	BnaC08g48540D	CDY52765.1	25.5	226	23–198	8.4	8
BnCYP25-3	BnaA09g29400D	CDY16206.1	25.6	226	20–200	6.5	8
BnCYP25-4	BnaA03g21710D	CDX83335.1	25.9	231	79–230	8.9	6
BnCYP25-5	BnaC03g25990D	CDX95718.1	25.9	231	79–230	8.9	6
BnCYP26-1	BnaA05g30740D	CDY24675.1	26.3	234	74–230	8.7	7
BnCYP26-2	BnaA08g08200D	CDX76580.1	26.3	239	151–213	6.5	4
BnCYP26-3	BnaC05g45190D	CDY05284.1	26.3	234	74–230	9.0	7

Continued

Name	Gene Name	GenBank Accession	kDa	aa	CLD	pI (theoretical)	Exons
BnCYP26-4	BnaA04g17840D	CDY29643.1	26.3	248	77–245	6.1	1
BnCYP26-5	BnaCnng32180D	CDY57615.1	26.5	236	76–232	8.7	7
BnCYP27-1	BnaC05g44950D	CDY05308.1	27.1	247	6–200	4.9	3
BnCYP27-2	BnaC08g31970D	CDX76611.1	27.4	253	81–249	8.7	7
BnCYP27-3	BnaC03g05790D	CDX70495.1	27.8	257	80–252	9.6	7
BnCYP27-4	BnaA09g39610D	CDY11395.1	27.9	258	86–254	9.0	7
BnCYP27-5	BnaA03g04200D	CDX78508.1	27.9	256	79–251	9.7	7
BnCYP28-1	BnaA04g27460D	CDY58761.1	28.1	258	86–254	8.8	7
BnCYP28-2	BnaC04g21530D	CDX93994.1	28.6	261	89–257	8.9	7
BnCYP30-1	BnaC08g08050D	CDY12400.1	30.4	280	68–267	6.3	2
BnCYP30-2	BnaA08g07170D	CDY41762.1	30.5	281	75–270	5.9	2
BnCYP31	BnaC07g00280D	CDY05775.1	32.0	277	59–173	4.8	7
BnCYP34-1	BnaC02g22620D	CDY45858.1	34.5	320	92–314	9.1	3
BnCYP34-2	BnaA02g16680D	CDY49138.1	34.5	320	91–314	9.1	3
BnCYP37-1	BnaC01g21230D	CDY35601.1	37.4	340	161–319	6.1	5
BnCYP37-2	BnaA01g17950D	CDY42959.1	37.5	340	161–316	6.6	5
BnCYP40-1	BnaC07g05260D	CDX71247.1	40.3	361	3–175	6.3	8
BnCYP40-2	BnaA07g04040D	CDY36653.1	40.3	361	3–175	6.0	8
BnCYP47-1	BnaC08g10870D	CDY43466.1	47.1	418	157–312	8.2	7
BnCYP47-2	BnaA05g33830D	CDY37518.1	47.2	433	254–433	5.1	7
BnCYP47-3	BnaC05g48850D	CDY49057.1	47.4	436	257–436	5.0	6
BnCYP49	BnaA05g24020D	CDX97677.1	49.8	459	273–438	6.6	12
BnCYP50	BnaC05g38110D	CDX98564.1	51.0	470	284–445	5.7	12
BnCYP52	BnaC01g05170D	CDX68936.1	52.4	467	6–186	8.6	11
BnCYP55	BnaA01g03810D	CDX75321.1	55.3	492	6–185	7.7	10
BnCYP62	BnaAnng12550D	CDY53714.1	62.0	556	4–177	10.7	13
BnCYP65-1	BnaA06g24990D	CDY08725.1	65.0	597	335–508	8.0	11
BnCYP65-2	BnaC03g48580D	CDY36499.1	65.1	597	335–508	7.3	11
BnCYP67	BnaC04g20680D	CDX93909.1	67.3	612	77–251	10.6	15
BnCYP70-1	BnaC03g54740D	CDY08708.1	70.1	622	465–619	6.0	13
BnCYP70-2	BnaA06g18830D	CDX99192.1	70.2	622	465–619	6.0	13
BnCYP86-1	BnaA01g04590D	CDX75243.1	86.9	765	3–174	11.8	13
BnCYP86-2	BnaC01g06080D	CDX68845.1	87.0	765	3–174	11.7	13
BnCYP112	BnaC03g71020D	CDY10512.1	112.6	992	1–176	5.9	18
BnCYP146	BnaA03g29520D	CDX74095.1	146.1	1268	1108–1264	6.0	13

**Table 1.** Nomenclature, gene name, accessions, molecular weight, protein sequence length, information about aligned amino acids, theoretical isoelectric point, and predicted exons of the *B. napus* CYP family.

weight. We used the molecular weight of the immature proteins as the basis, because the prediction of potential cleavable signal peptides by different prediction tools was not unambiguous. Table 1 summarizes the information about all identified BnCYPs.

**The cyclophilin-like domain.** Since CYPs are characterized by the highly conserved CLD, its occurrence in the identified potential CYP sequences was verified. Most of the BnCYPs contained full length CLDs, whereas CYPs with a molecular weight below 17 kDa have partial CLDs missing essential residues or whole secondary structure parts and thereby might be lacking PPIase activity.

As a consequence of the CLD analysis, two proteins annotated as CYPs in the first place, BnaC05g10020D (30 kDa) and BnaC04g45810D (54 kDa), were excluded from subsequent studies. In addition, for one of the putative low molecular weight CYPs, BnCYP7-1, no CLD was predicted. Nevertheless, it was retained, because a BLASTp search and sequence alignment suggested a partial CLD.

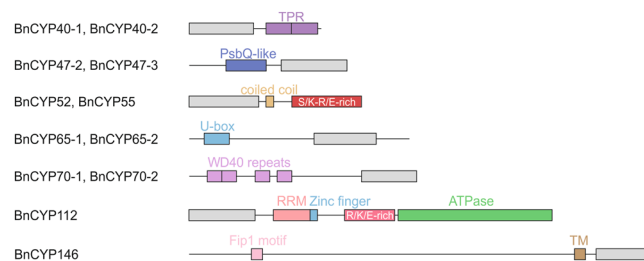
Figure 1 depicts the conserved sequence of all identified *B. napus* CLDs aligned with the secondary structure of human cyclophilin A (hCYPA) as a reference (for a detailed view of the multiple alignment see Supplementary Fig. S1). The crystal structure of hCYPA shows eight  $\beta$ -strands forming a  $\beta$ -barrel structure, plus two additional  $\alpha$ -helices located at the top and bottom<sup>52</sup>. Human CYPA is often referred to as the “archetypal” CYP. Residues important for CYP function as well as residues promoting the secondary structure are conserved, whereas gaps in the conserved sequence represent insertions in individual members of this protein family. In previous studies it has been demonstrated that the highly conserved amino acid W121 of hCYPA (Fig. 1, W225) is required for CsA binding and interacts directly with the phosphatase calcineurin, but it is not essential for PPIase activity<sup>53,54</sup>.

AtCYP	Gene	Subcellular Localization*	Comments	BnCYP homolog (Identity)
AtCYP18-1	At1g01940	Cytosol (p <sup>47,48</sup> )	up-regulated during heat stress <sup>81</sup>	BnCYP21-1 (98%)
AtCYP18-2	At2g36130	Cytosol (p <sup>47,48</sup> )	pre-mRNA splicing, down-regulated after pathogen treatment <sup>48,55</sup>	BnCYP18-1 (95%), BnCYP13 (79%), BnCYP8 (56%)
AtCYP18-3	At4g38740	Cytosol (p <sup>47,48</sup> )	plant growth (stem elongation, shoot branching), hormone signalling (links light signal receptors to brassinosteroids), pathogen defence/ETI (by binding of pathogenic proteins/RNA (inhibition of replication (TBSV), interacts with <i>A. tumefaciens</i> VirD2 protein, <i>P. syringae</i> AvrPt2 protease and the plant RIN4 protein)), up-regulated after salt treatment <sup>48,82-89</sup>	BnCYP18-4 (93%), BnCYP18-6 (92%), BnCYP18-7 (92%)
AtCYP18-4	At4g34870	Cytosol (p <sup>47,48</sup> )	pathogen infection (interaction with <i>A. tumefaciens</i> VirD2 protein), down-regulated after salt and cytokinin treatment <sup>48,85</sup>	BnCYP18-2 (83%), BnCYP18-3 (83%), BnCYP17-1 (74%), BnCYP10-1 (35%)
AtCYP19-1	At2g16600	Cytosol (p <sup>47,48</sup> )	seed development, pathogen defence (ROS production, inhibiting pathogen growth), up-regulated after <i>P. syringae</i> infection, down-regulated after ABA treatment <sup>40,48,90</sup>	BnCYP18-5 (92%), BnCYP18-8 (92%), BnCYP26-4 (66%), BnCYP22-3 (65%), BnCYP20 (62%), BnCYP14-1 (62%), BnCYP14-2 (62%), BnCYP14-3 (60%), BnCYP14-6 (52%), BnCYP16 (50%), BnCYP7-1 (36%)
AtCYP19-2	At2g21130	Cytosol (p <sup>47,48</sup> )	—	BnCYP7-3 (59%), BnCYP27-1 (46%)
AtCYP19-3	At3g56070	Cytosol (p <sup>47,48</sup> )	pathogen defence (inhibition of replication by binding to viral replication proteins (TBSV)), Ca <sup>2+</sup> signalling, interaction with calmodulin (35–70 aa), sensitive to Cu <sup>2+</sup> (implying redox regulation) <sup>84,91</sup>	BnCYP19-1 (91%), BnCYP19-2 (90%), BnCYP5 (48%)
AtCYP19-4	At2g29960	Cytosol+SP (e <sup>92,93</sup> )	cell polarity (regulates GNOM activity during embryogenesis), up-regulated after salt and cytokinin treatment <sup>48,92,93</sup>	BnCYP21-5 (93%), BnCYP21-3 (92%), BnCYP12-1 (56%), BnCYP14-4 (52%)
AtCYP20-1	At5g58710	SP (p <sup>47,48</sup> )	UPR (unfolded protein response) in the ER, up-regulated upon ER stress, interaction with PP2A a component of multiple signalling pathways (e.g. auxin transport, growth response) <sup>94,95</sup>	BnCYP21-9 (97%), BnCYP21-4 (92%), BnCYP21-6 (92%)
AtCYP20-2	At5g13120	TL (e <sup>96,97</sup> )	down-regulated after pathogen treatment and up-regulated after light treatment, NAD(P)H dehydrogenase complex subunit, PPLase activity induced by oxidative stress, plant development (interacting with BZR1, a transcription factor responding to brassinosteroids (hormone signalling)), general protein folding catalyst in the TL <sup>48,62,63,97-100</sup>	BnCYP27-3 (86%), BnCYP27-5 (86%), BnCYP7-2 (51%)
AtCYP20-3	At3g62030	Stroma (e <sup>96,97,101</sup> )	light and oxidative stress, redox regulation suggested (interacts with peroxiredoxins PrxA and PrxB), cysteine biosynthesis (interaction with SAT1 (serine acetyltransferase)), JA signalling, binds to JA and OPDA, interacts with <i>A. tumefaciens</i> VirD2 protein <sup>38,40,85,102-104</sup>	BnCYP27-2 (89%), BnCYP27-4 (89%), BnCYP28-1 (85%), BnCYP28-2 (85%), BnCYP14-5 (81%), BnCYP24-3 (79%)
AtCYP21-1	At4g34960	SP (p <sup>47,48</sup> )	—	BnCYP24-5 (95%), BnCYP24-6 (95%), BnCYP24-4 (85%), BnCYP24-7 (83%)
AtCYP21-2	At3g55920	SP (p <sup>47,48</sup> )	water stress <sup>105</sup>	BnCYP24-1 (92%), BnCYP24-2 (92%)
AtCYP21-3	At2g47320	Mitochondria (p <sup>47,48</sup> )	—	BnCYP25-4 (83%), BnCYP25-5 (83%)
AtCYP21-4	At3g66654	Mitochondria (p <sup>47,48</sup> )	down-regulated after dark treatment <sup>48</sup>	BnCYP26-3 (89%), BnCYP26-1 (88%), BnCYP146 (87%), BnCYP26-5 (86%)
AtCYP22	At2g38730	Cytosol (p <sup>47,48</sup> )	—	BnCYP21-2 (93%), BnCYP21-7 (93%), BnCYP21-8 (92%), BnCYP22-2 (92%)
AtCYP23	At1g26940	SP (p <sup>47,48</sup> )	—	BnCYP25-3 (92%), BnCYP25-2 (87%), BnCYP25-1 (86%)
AtCYP26-1	At3g22920	Cytosol (p <sup>47,48</sup> )	—	—
AtCYP26-2	At1g74070	TL (p <sup>47,48,106</sup> )	down-regulated after sucrose treatment <sup>48</sup>	BnCYP34-1 (84%), BnCYP34-2 (83%)
AtCYP28	At5g35100	TL (e <sup>96,106</sup> )	down-regulated after dark and high CO <sub>2</sub> treatment <sup>48</sup>	BnCYP30-1 (86%), BnCYP30-2 (85%)
AtCYP37	At3g15520	TL (e <sup>96,106</sup> )	down-regulated after dark treatment <sup>48</sup>	BnCYP49 (87%), BnCYP50 (87%)

Continued

AtCYP	Gene	Subcellular Localization*	Comments	BnCYP homolog (Identity)
AtCYP38	At3g01480	TL (e <sup>96,97</sup> )	no PPIase activity in the TL, PsbQ-like domain, photo system II: folding of subunits and assembly, down-regulated after dark treatment and up-regulated after light treatment <sup>48,61-65</sup>	BnCYP47-2 (86%), BnCYP47-3 (84%), BnCYP22-1 (69%), BnCYP31 (51%)
AtCYP40	At2g15790	Cytosol (p <sup>47,48</sup> )	interaction with HSP90, AGO1, miRNA156, regulating RISC complex, regulation of vegetative phase change <sup>56,57,107</sup>	BnCYP40-2 (93%), BnCYP40-1 (92%)
AtCYP57	At4g33060	Cytosol/Nucleus (p <sup>47,48</sup> )	pathogen defence (callose accumulation), up-regulated after <i>P. syringae</i> infection, RNA-interacting region <sup>30,47,48</sup>	BnCYP55 (86%), BnCYP52 (80%)
AtCYP59	At1g53720	Nucleus (e <sup>67</sup> )	cyclophilin-RNA interacting protein (CRIP), Zinc finger motif, pre-mRNA processing, transcription (modulates RNA polymerase II activity) <sup>32,36,67</sup>	BnCYP12-2 (97%), BnCYP112 (85%), BnCYP10-2 (69%)
AtCYP63	At3g63400	Nucleus (p <sup>47,48</sup> )	may be involved in RNA metabolism <sup>47</sup>	BnCYP62 (65%), BnCYP67 (65%)
AtCYP65	At5g67530	Cytosol (p <sup>47,48</sup> )	suggested to be involved in the ubiquitin degradation pathway <sup>47</sup>	BnCYP65-1 (92%), BnCYP65-2 (92%), BnCYP26-2 (33%), BnCYP47-1 (27%), BnCYP37-1 (26%), BnCYP37-2 (26%)
AtCYP71	At3g44600	Nucleus (e <sup>35</sup> )	plant development, gene expression (histone modification, chromatin assembly), down-regulated after auxin treatment and after knox (transcription factors) overexpression <sup>35,48,60</sup>	BnCYP70-1 (91%), BnCYP70-2 (91%)
AtCYP95	At4g32420	Nucleus (p <sup>47,48</sup> )	may be involved in RNA metabolism <sup>47</sup>	BnCYP86-1 (69%), BnCYP86-2 (67%)

**Table 2.** The Arabidopsis cyclophilin family, localization, functions and corresponding *Brassica* homologs. \*Localization was either p = predicted (SP, secretory pathway; TL, thylakoid lumen) or e = experimentally proven as described in the references.

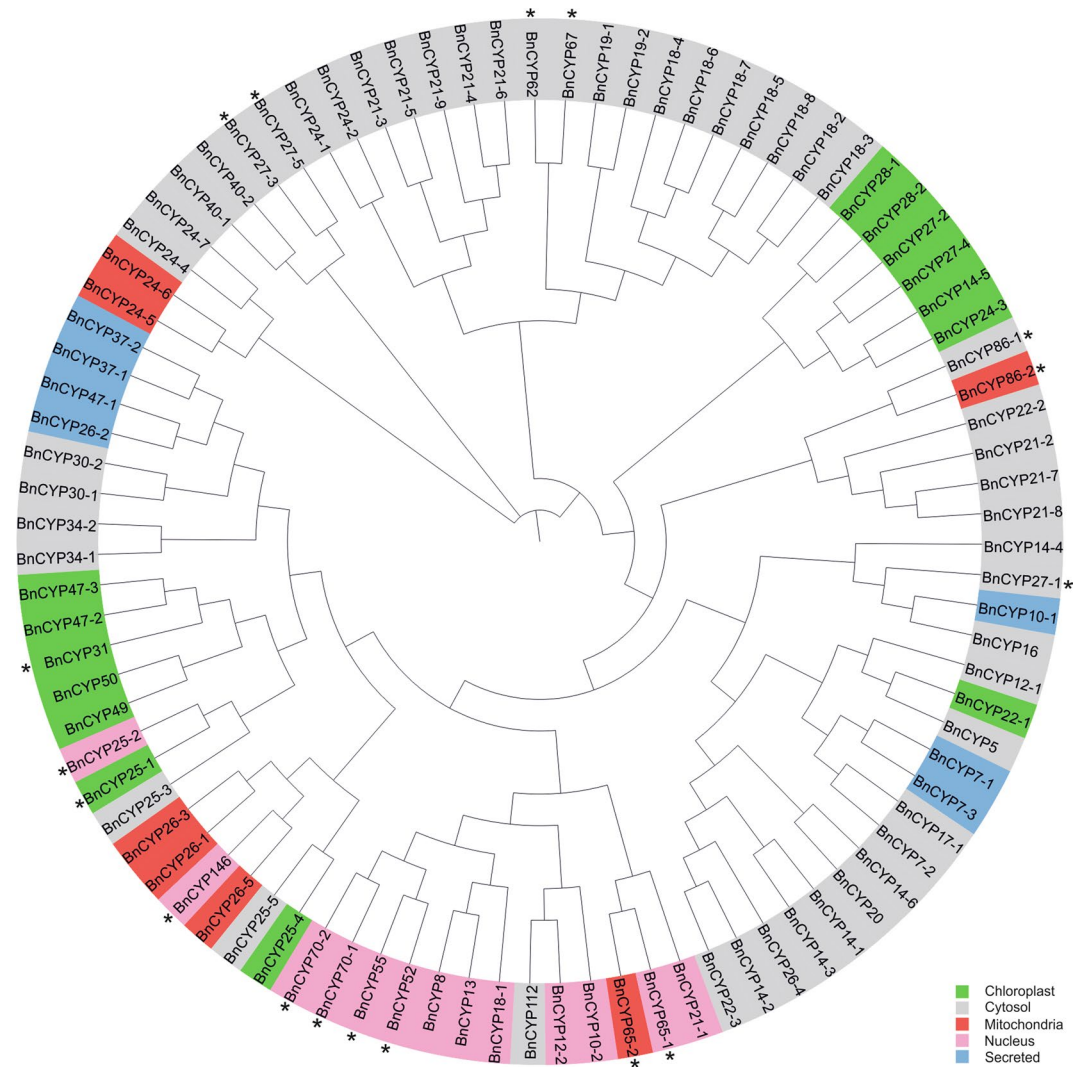


**Figure 2.** Domain structure of multi-domain *B. napus* cyclophilins. While most of the CYPs are single-domain proteins, 12 BnCYPs possess additional domains for specific tasks. CLDs are represented by a grey box. RRM = RNA recognition motif, TPR = tetratricopeptide repeat, TM = transmembrane domain.

This amino acid was not present in 57 out of the 91 BnCYP proteins. Further highly conserved amino acids fundamental for the PPIase activity of hCYP are R55 (Fig. 1, R87), F60 (Fig. 1, F102;), and H126 (Fig. 1, H232)<sup>53</sup>. Three single mutations of these amino acids reduced the original activity of the human wild-type isomerase to less than 1%. However, the mutant proteins were still able to bind CsA. Since some of the BnCYPs did not possess all of these three amino acids, their PPIase activity needs to be experimentally confirmed. The motif VXGXV of the hydrophobic region around  $\beta$ -sheet-VII is reported to be conserved among all AtCYP proteins<sup>47</sup>. This motif was also present in 62 BnCYP protein sequences (Fig. 1, V234, G236, V238).

Exceptional insertions of amino acids in plant CYPs between  $\beta$ -sheet-I and  $\beta$ -sheet-II ( $\beta$ -I/ $\beta$ -II),  $\beta$ -sheet-IV and  $\beta$ -sheet-V ( $\beta$ -IV/ $\beta$ -V), and  $\beta$ -sheet-VI and  $\alpha$ -helix-II ( $\beta$ -VI/ $\alpha$ -II)<sup>20</sup> have also been observed in this study. In more detail, 3 amino acids are inserted between  $\beta$ -I/ $\beta$ -II (Fig. 1, residues 25–27) in 16 BnCYPs. Yet, 45 BnCYPs do not show this insertion. Interestingly, we also observed deletions in 30 BnCYPs in this region. Only a minority of BnCYPs contains an insertion between  $\beta$ -IV/ $\beta$ -V (Fig. 1, residues 133–152) with 1 to 19 amino acids in length, but also here deletions occurred in some cases. Insertions between  $\beta$ -VI/ $\alpha$ -II (Fig. 1, residues 216–221) are only present in 5 BnCYPs with medium and higher molecular masses. An additional insertion between  $\alpha$ -helix-I and  $\beta$ -sheet-III ( $\alpha$ -I/ $\beta$ -III junction) has been described by Romano *et al.*<sup>47</sup>. Here, 8 to 11 amino acids are inserted in several AtCYPs and 3 to 4 amino acids in chloroplast variants. For *B. napus* a 7 amino acid insertion occurs in several CYPs. For example, except for BnCYP18-1, all 18 and 19 kDa BnCYPs show this insertion.

**Domain architecture of cyclophilins and homology.** Depending on the domains present, CYPs are classified as single- (SD) and multi-domain (MD) forms<sup>48</sup>. SD CYPs contain only a CLD which was the case for most of the BnCYPs with a rather low molecular weight (18–30 kDa). While there are mostly two BnCYP homologs with high sequence identity (>80%) corresponding to each AtCYP, caused by polyploidization, there are only a few exceptions with only one corresponding BnCYP homolog. One example is BnCYP18-1 which is



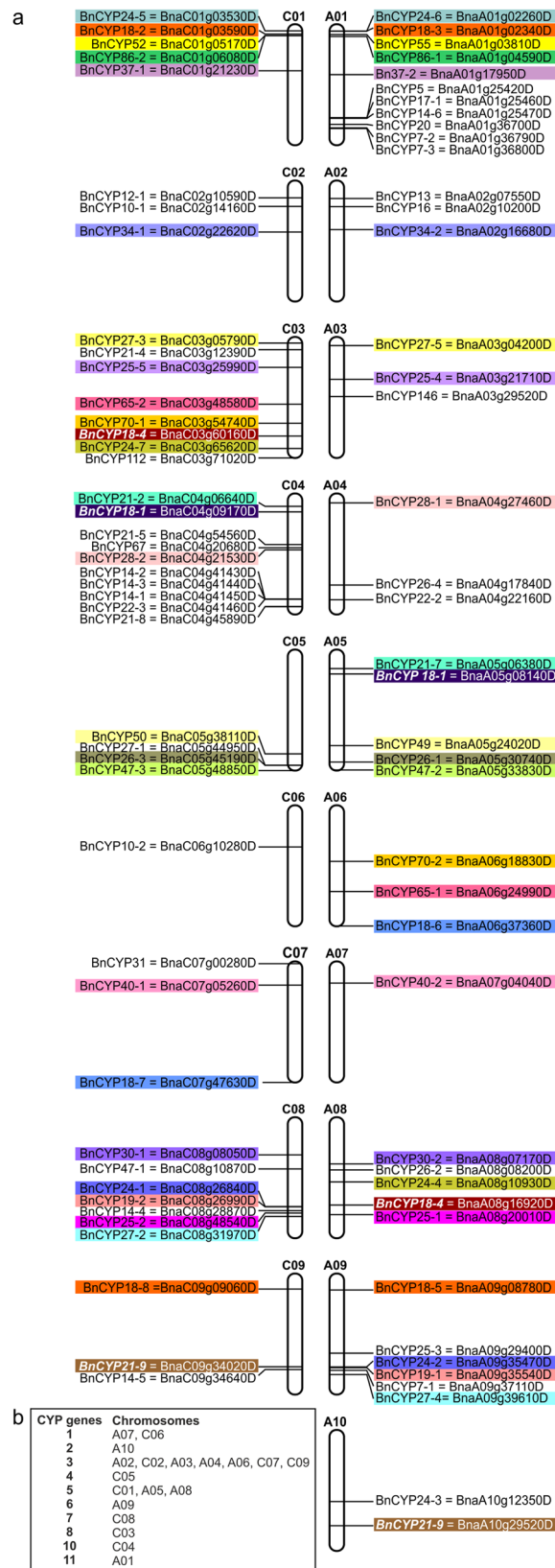
**Figure 3.** Phylogenetic tree of *B. napus* cyclophilins. The phylogenetic analysis was based on a sequence alignment by ClustalOmega and constructed as described in Materials and Methods with ignored branch lengths. As indicated by colour code, this protein family is distributed over all intercellular organelles. In addition, some BnCYPs possess putative nuclear localization signals (marked by \*).

the only *B. napus* homolog of AtCYP18-2 with a full length CLD (with 95% sequence identity, Table 2). It has been shown that the Arabidopsis CYP AtCYP18-2 is recruited by AtSKIP to the nucleus to regulate pre-mRNA splicing<sup>55</sup>. Because of the high homology, the oilseed rape variant might fulfil a similar role in the nucleus of *B. napus* cells.

For 12 *B. napus* CYPs additional domains are predicted which promote diverse capabilities like interaction (protein-protein, protein-DNA, protein-RNA) and modification (ubiquitination) (Fig. 2). These seem to be conserved among *Brassicaceae*, since similar domain structures exist in the corresponding *A. thaliana* homologs.

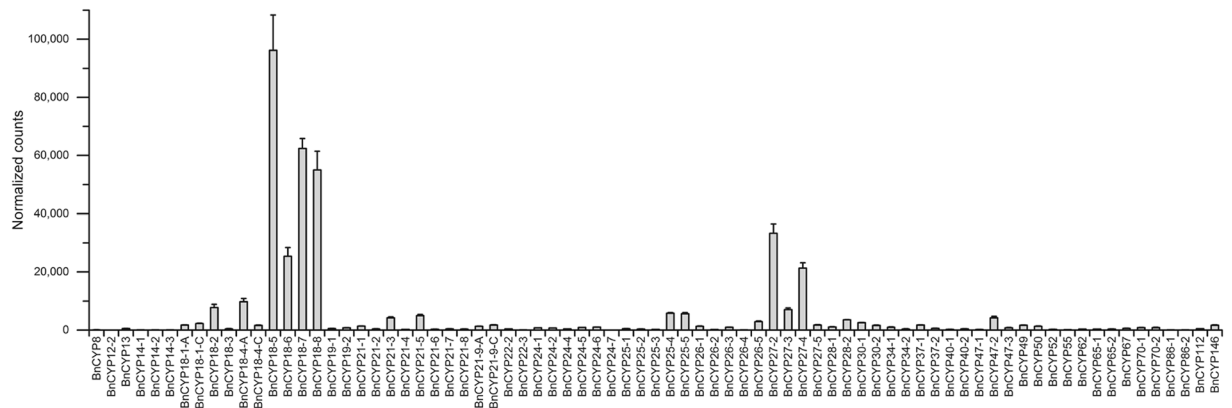
The MD CYPs BnCYP40-1 and BnCYP40-2 are characterized by a tetratricopeptide repeat (TPR) motif with two TPRs, each 34 amino acids long. Such motifs mediate protein-protein interactions and can thereby assist in the assembly of multi-protein complexes. BnCYP40-1 and BnCYP40-2 show a high degree of sequence identity (92 and 93%, respectively) to AtCYP40 (Table 2). A cytoplasmic interaction partner of AtCYP40 is HSP90, which mediates its recruitment to an intermediate RISC complex by its TPR motifs<sup>56,57</sup>. Thereby, AtCYP40 forms, together with HSP90, a complex with AGO1 and a small RNA duplex before the mature RISC complex consisting of AGO1 and a siRNA or miRNA strand is formed<sup>56</sup>. The mechanism of CYP40-HSP90 binding is conserved between different species. Besides plants, human HSP90 is also known to bind to hCYP40, the human AtCYP40 homolog<sup>58</sup>. Thus, a similar function for the highly identical CYPs BnCYP40-1 and BnCYP40-2 can be assumed.

Other CYPs showing classical protein-protein interaction domains are BnCYP70-1 and BnCYP70-2, both containing four tryptophan-aspartic acid (WD40) repeats. These are short structural motifs typically forming a four stranded anti-parallel  $\beta$ -sheet. Multiple copies build a circular  $\beta$ -propeller structure promoting protein-protein interactions. Both show 91% sequence identity to their homolog in Arabidopsis, AtCYP71 (Table 2), which is located in the nucleus and functions in the regulation of gene expression and organogenesis.

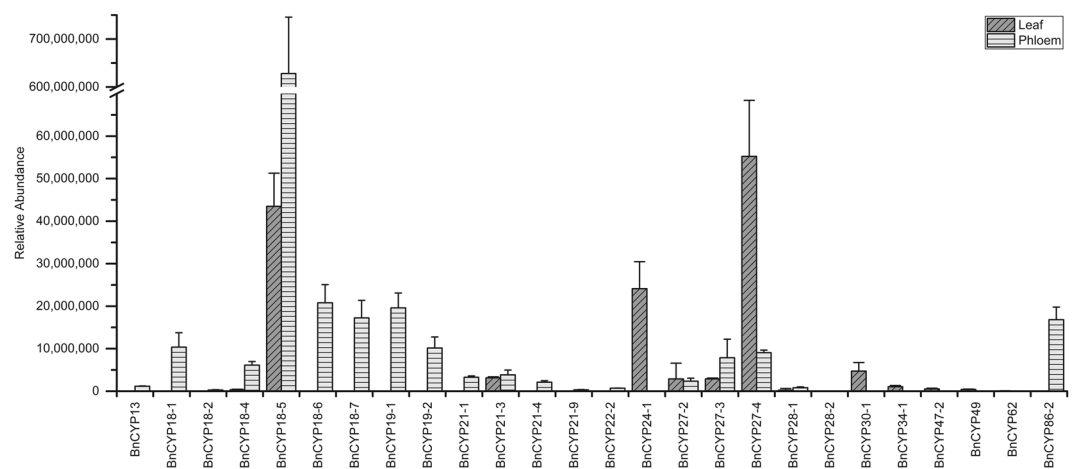


**Figure 4.** Genomic distribution of *cyclophilin* genes on *B. napus* chromosomes. (a) Chromosomal locations of *BnCYP* genes are indicated based on the information provided by the *Brassica napus* Genome Browser. Highlighted by identical colours are *BnCYP* genes which encode protein isoforms. Genes encoding identical proteins are written in italics and white colour. (b) Summary of the number of *BnCYP* genes encoded on each chromosome.





**Figure 5.** Expression pattern of *BnCYP* genes from leaf material. In case of two genes encoding the same BnCYP protein, the transcript is abbreviated by its BnCYP name with -A or -C depending on the chromosomal location of the gene, respectively. Raw data are provided in Supplementary Table S2.



**Figure 6.** Relative abundance of leaf and phloem BnCYP proteins. The relative abundance of each BnCYP refers to the abundance of its unique peptides either from leaf extract ( $n = 3$ ) or phloem sap ( $n = 4$ ) samples.

The WD40 domain enables AtCYP71 to associate with histone H3 affecting its methylation<sup>35</sup>. Data from yeast PPIases provide evidence that they are responsible for histone modification by utilizing the peptidyl-prolyl bond isomerisation as a molecular switch for transcription regulation<sup>59</sup>. Furthermore, AtCYP71 was shown to interact with FAS1 (a subunit of Chromatin Assembly Factor-1) and LHP1 (a heterochromatin protein)<sup>60</sup>, suggesting that AtCYP71 is involved in histone modification and chromatin assembly.

BnCYP47-2 and BnCYP47-3 both contain a putative PsbQ-like domain and have only small differences in their N-terminal amino acid sequences. Their closest homolog is AtCYP38 with 86 and 84% sequence identity (Table 2), which also contains a PsbQ-like domain. This domain is typical for proteins localized in the chloroplast, but its function is mostly unknown. AtCYP38 is the only Arabidopsis CYP for which a crystal structure is available. This revealed the additional PsbQ-like domain<sup>61</sup>. AtCYP38 is experimentally assigned to the chloroplast thylakoid lumen, does not show any PPIase activity<sup>62,63</sup>, and plays a critical role in the assembly and maintenance of photosystem II<sup>61,64,65</sup>. It is suggested to interact with the E-loop of chlorophyll protein47 (CP47), a component of the photosystem II (PSII) complex<sup>61</sup>, via its CLD. Furthermore it is proposed to be responsible for proper folding and insertion of D1 and CP43, both components of PSII<sup>65</sup>.

There are also two BnCYPs showing classical domains for protein modification. BnCYP65-1 and BnCYP65-2 both contain a Zinc finger U-box motif. Therefore, they might be involved in the ubiquitin degradation pathway.

Four BnCYPs contain sequences characteristic for RNA-interacting proteins. BnCYP52 and BnCYP55 both possess a coiled coil domain. This domain represents a structural motif with coiled  $\alpha$ -helices. Proteins containing such a domain can, for example, be transcription factors involved in the regulation of gene expression. The respective *A. thaliana* homolog is AtCYP57, which shows, besides a predicted coiled coil domain, also an S/K-R/E-rich region<sup>47</sup>. Sequence comparisons of BnCYP52, BnCYP55 and AtCYP57 revealed a high sequence identity in the AtCYP57 S/K-R/E-rich region (data not shown). This region is supposed to mediate the interaction with ribonucleoproteins<sup>66</sup>.

Name	Gene Name	BnCYPs identified by RNA-Seq	BnCYPs identified by LC-MS/MS	
			Leaf	Phloem
BnCYP5	BnaA01g25420D			
BnCYP7-1	BnaA09g37110D			
BnCYP7-2	BnaA01g36790D			
BnCYP7-3	BnaA01g36800D			
BnCYP8	BnaCnng42430D	x		
BnCYP10-1	BnaC02g14160D			
BnCYP10-2	BnaC06g10280D			
BnCYP12-1	BnaC02g10590D			
BnCYP12-2	BnaAnng41240D	x		
BnCYP13	BnaA02g07550D	x		x
BnCYP14-1	BnaC04g41450D	x		
BnCYP14-2	BnaC04g41430D	x		
BnCYP14-3	BnaC04g41440D	x		
BnCYP14-4	BnaC08g28870D			
BnCYP14-5	BnaC09g34640D			
BnCYP14-6	BnaA01g25470D			
BnCYP16	BnaA02g10200D			
BnCYP17-1	BnaA01g25460D			
BnCYP18-1	BnaA05g08140D	x		x
	BnaC04g09170D	x		
BnCYP18-2	BnaC01g03590D	x		x
BnCYP18-3	BnaA01g02340D	x		
BnCYP18-4	BnaC03g60160D	x	x	x
	BnaA08g16920D	x		
BnCYP18-5	BnaA09g08780D	x	x	x
BnCYP18-6	BnaA06g37360D	x		x
BnCYP18-7	BnaC07g47630D	x		x
BnCYP18-8	BnaC09g09060D	x		
BnCYP19-1	BnaA09g35540D	x		x
BnCYP19-2	BnaC08g26990D	x	x	x
BnCYP20	BnaA01g36700D			
BnCYP21-1	BnaCnng08980D	x		x
BnCYP21-2	BnaC04g06640D	x		
BnCYP21-3	BnaAnng15590D	x	x	x
BnCYP21-4	BnaC03g12390D	x		x
BnCYP21-5	BnaC04g54560D	x		
BnCYP21-6	BnaAnng17350D	x		
BnCYP21-7	BnaA05g06380D	x		
BnCYP21-8	BnaC04g45890D	x		
BnCYP21-9	BnaA10g29520D	x		x
	BnaC09g34020D	x		
BnCYP22-1	BnaCnng32070D			
BnCYP22-2	BnaA04g22160D	x		x
BnCYP22-3	BnaC04g41460D	x		
BnCYP24-1	BnaC08g26840D	x	x	
BnCYP24-2	BnaA09g35470D	x		
BnCYP24-3	BnaA10g12350D			
BnCYP24-4	BnaA08g10930D	x		
BnCYP24-5	BnaC01g03530D	x		
BnCYP24-6	BnaA01g02260D	x		
BnCYP24-7	BnaC03g65620D	x		
BnCYP25-1	BnaA08g20010D	x		
BnCYP25-2	BnaC08g48540D	x		
BnCYP25-3	BnaA09g29400D	x		
Continued				

Name	Gene Name	BnCYPs identified by RNA-Seq	BnCYPs identified by LC-MS/MS	
			Leaf	Phloem
BnCYP25-4	BnaA03g21710D	x		
BnCYP25-5	BnaC03g25990D	x		
BnCYP26-1	BnaA05g30740D	x		
BnCYP26-2	BnaA08g08200D	x		
BnCYP26-3	BnaC05g45190D	x		
BnCYP26-4	BnaA04g17840D	x		
BnCYP26-5	BnaCnng32180D	x		
BnCYP27-1	BnaC05g44950D			
BnCYP27-2	BnaC08g31970D	x	x	x
BnCYP27-3	BnaC03g05790D	x	x	x
BnCYP27-4	BnaA09g39610D	x	x	x
BnCYP27-5	BnaA03g04200D	x		
BnCYP28-1	BnaA04g27460D	x	x	x
BnCYP28-2	BnaC04g21530D	x		x
BnCYP30-1	BnaC08g08050D	x	x	
BnCYP30-2	BnaA08g07170D	x		
BnCYP31	BnaC07g00280D			
BnCYP34-1	BnaC02g22620D	x	x	
BnCYP34-2	BnaA02g16680D	x		
BnCYP37-1	BnaC01g21230D	x		
BnCYP37-2	BnaA01g17950D	x		
BnCYP40-1	BnaC07g05260D	x		
BnCYP40-2	BnaA07g04040D	x		
BnCYP47-1	BnaC08g10870D	x		
BnCYP47-2	BnaA05g33830D	x	x	
BnCYP47-3	BnaC05g48850D	x		
BnCYP49	BnaA05g24020D	x	x	
BnCYP50	BnaC05g38110D	x		
BnCYP52	BnaC01g05170D	x		
BnCYP55	BnaA01g03810D	x		
BnCYP62	BnaAnng12550D	x	x	
BnCYP65-1	BnaA06g24990D	x		
BnCYP65-2	BnaC03g48580D	x		
BnCYP67	BnaC04g20680D	x		
BnCYP70-1	BnaC03g54740D	x		
BnCYP70-2	BnaA06g18830D	x		
BnCYP86-1	BnaA01g04590D	x		
BnCYP86-2	BnaC01g06080D	x		x
BnCYP112	BnaC03g71020D	x		
BnCYP146	BnaA03g29520D	x		

**Table 3.** Identified *B. napus* cyclophilins by RNA-Seq and LC-MS/MS.

BnCYP112 contains an RNA recognition motif (RRM), a Zinc finger motif (CCHC-type), a positively charged region (arginine, lysine, glutamate) and an actin-like ATPase domain. Besides AtCYP59 in *A. thaliana*, homologs to BnCYP112 can be found in different species like *P. tetraurelia*, *S. pombe*, *C. elegans*, *D. melanogaster*, and *H. sapiens*<sup>32</sup>. This group of proteins is called CRIPs, for cyclophilin-RNA interacting proteins<sup>32</sup>. Characteristic for this group of proteins is an RRM in addition to the CLD. Nevertheless, CRIPs differ in their C-terminal region. AtCYP59 and BnCYP112 are the only members containing a Zinc finger motif. AtCYP59 interacts with the C-terminal domain of the largest subunit of RNA polymerase II and thereby influences its phosphorylation state. Furthermore, binding of an RNA transcript decreases PPIase activity of AtCYP59. This might modulate RNA polymerase II activity<sup>36</sup>. Therefore, it is suggested that AtCYP59 connects pre-mRNA processing and transcription<sup>36, 67</sup>. Moreover, BnCYP112 is the only homolog of all CRIPs showing an additional actin-like ATPase domain.

BnCYP146 is the largest CYP found in oilseed rape. It contains a transmembrane domain and a Fip1 motif. In yeast, the Fip1 protein ensures the polyadenylation of mRNAs by interacting with poly(A) polymerase. There is no *A. thaliana* homolog showing the same domain structure as BnCYP146, but its CLD is closely related to AtCYP21-4 (Table 2).

**Subcellular localization and phylogenetic relationships.** The phylogenetic analysis of the full length amino acid sequences revealed clusters of proteins with high sequence similarity (Fig. 3). Often, these groups possess the same additional domains, consistent subcellular localization, and the same *A. thaliana* CYP homolog.

As previously described for other organisms (e.g. *A. thaliana* homologs, see Table 2), *B. napus* CYPs are targeted to all intercellular organelles. At least 50 BnCYPs are predicted to be located in the cytosol. This result is comparable to the subcellular distribution of the *A. thaliana* CYPs where 14 of 29 are either predicted or experimentally proven to be located in the cytosol (Table 2). Furthermore, 14 BnCYPs are predicted to be located in chloroplasts, 7 in mitochondria, 13 in the nucleus, and 7 are secreted (summarized in Fig. 3, for a detailed prediction see Supplementary Table S1). In addition, some BnCYPs contain nuclear localization signals (NLS). Proteins predicted to be cytosolic or nuclear which possess NLS might relocate between both compartments.

**Genomic distribution of cyclophilins.** *CYP* genes are distributed on all 19 chromosomes of the *B. napus* genome (Fig. 4a). There is no chromosome that is not encoding any *CYP*. Chromosomes A07 and C06 contain only one *CYP* each. Chromosome A01 has the largest number with 11 *CYP* genes. All other chromosomes contain between 2 and 10 *CYP* genes (Fig. 4b). The 4 *CYP* genes *BnaAnng41240D* (*BnCYP12-2*), *BnaAnng15590D* (*BnCYP21-3*), *BnaAnng17350D* (*BnCYP21-6*), and *BnaAnng12550D* (*BnCYP62*) were mapped to the A chromosomes, but without a detailed information regarding the exact chromosomal location. Likewise, the 4 *CYP* genes *BnaCnng42430D* (*BnCYP8*), *BnaCnng08980D* (*BnCYP21-1*), *BnaCnng32070D* (*BnCYP22-1*), and *BnaCnng32180D* (*BnCYP26-5*) were mapped to the C chromosomes, but again without a detailed information regarding the exact chromosomal location. The A chromosomes contain 42 *CYP* genes and 4 not fully assigned genes, and the C chromosomes contain 44 *CYP* genes and 4 not fully assigned genes. Thus, both chromosome sets (A and C), which originate from the *B. rapa* and *B. oleracea* genome, contain a similar number of *CYP* genes.

Interestingly, there are three *CYP* proteins that are each encoded by two of the six following genes *BnaA05g08140D* and *BnaC04g09170D* (*BnCYP18-1*), *BnaA08g16920D* and *BnaC03g60160D* (*BnCYP18-4*), or *BnaA10g29520D* and *BnaC09g34020D* (*BnCYP21-9*). Thus, one copy of these genes occurs on each of the chromosomal sets. Alignments of these *CYP* genes from the A and respective C chromosome showed slightly different nucleotide sequences. Nevertheless, these encode identical amino acid sequences (data not shown).

Due to the two progenitor chromosomal sets, several *CYP* genes encode proteins with high sequence homologies. Most often these occur as pairs, each gene originating from one progenitor chromosomal set. They are highlighted with identical colours on the A and C chromosomes in Fig. 4 and are from now on called isoforms. Some of them are located on the respective A and C chromosome, others are spread on different chromosomes, probably due to chromosomal rearrangements of the *B. napus* genome (as recently described by Cheng *et al.*<sup>68</sup>).

In summary, the allopolyploidy of oilseed rape results in two isoforms of many proteins. These show high sequence identities on the nucleotide and amino acid levels and originate from either the A or C genome. Due to their similarity, they cluster in the phylogenetic tree and often the same localization is predicted (Fig. 3). They might either possess the ability to replace each other or might be specialized for certain tasks and interaction partners.

One of the biggest groups of isoforms among the BnCYPs is the group of 18 kDa proteins. These contain a full-length CLD with a high degree of conservation, a cytosolic localization, and can be further subdivided by their homology. The phylogenetic tree reveals isoform pairs: BnCYP18-2 and BnCYP18-3, BnCYP18-5 and BnCYP18-8, and BnCYP18-6 and BnCYP18-7 (Fig. 3). Each pair has a representative gene on chromosome A and C, respectively (Fig. 4a). Interestingly, BnCYP18-4 which is closely related to the pair BnCYP18-6/18-7 is encoded by two genes located on C03 and A08 that might derive from a gene duplication event. The *A. thaliana* homologs of the described 18 kDa isoforms are AtCYP18-3, AtCYP18-4, and AtCYP19-1 which are shown to be important for various processes from plant growth to pathogen defence (summarized in Table 2). In contrast to the other 18 kDa BnCYPs, BnCYP18-1 is the only 18 kDa CYP with introns and without the prominent insertion between  $\alpha$ -helix-I and  $\beta$ -sheet-III. Moreover, its localization is predicted to be nuclear instead of cytosolic (Fig. 3).

**Expression analysis of the predicted BnCYPs on transcript and protein levels.** To investigate the mRNA expression of the bioinformatically predicted *CYP* genes, RNA-Seq of RNA isolated from leaves was performed. This approach revealed 77 expressed *BnCYPs* under the applied conditions (summarized in Table 3). The expression pattern indicates a large variance in the abundance of *BnCYP* transcripts (Fig. 5, for raw data see Supplementary Table S2). The results thus show that most of the predicted *CYP* genes are indeed transcribed. Transcripts for *CYP* genes not detected in this study could be expressed in other tissues, specific cell types, or under different experimental conditions.

In the analysed leaf sample, the 18 kDa and 27 kDa family members belong to the strongest expressed ones, with *BnCYP18-5* showing the highest read count. Interestingly, some isoform pairs had differing read counts, e.g. *BnCYP18-2/BnCYP18-3*, *BnCYP18-4/BnCYP18-6/BnCYP18-7*, *BnCYP27-3/BnCYP27-5*.

Since CYPs are a prominent protein class in the phloem long-distance transport system of higher plants including oilseed rape with a potential function as a molecular chaperone important for protein long-distance transport<sup>2,41</sup>, we performed LC-MS/MS analysis of phloem protein extract and compared the CYP protein profile to that of leaf extract. Table 3 summarizes the BnCYPs identified from leaf and phloem protein extracts. In total, 26 BnCYPs could be detected at the protein level under the applied conditions with six being unique to leaves and 12 to phloem exudate. Since sieve elements do not have a functional transcription machinery, RNA-Seq was not performed with phloem samples. However, phloem samples contain a specific set of mobile RNAs that have in part been implicated with long-distance signalling<sup>43</sup>, but this was not the subject of the present study.

Most of the identified BnCYP proteins are predicted to be localized in the cytosol, but some potentially chloroplastic, mitochondrial and nuclear BnCYPs were found as well. Besides several SD CYPs, also one MD CYP,

the putative chloroplastic protein BnCYP47-2, could be identified in leaf extract. The analysis of phloem sap revealed a low molecular weight CYP, BnCYP13, which is suggested to contain only a partial CLD. Figure 6 shows the abundance of the identified CYP proteins in the two examined organs (more details in Supplementary Tables S3, S4, S5, S6). Whereas BnCYP18-5, BnCYP24-1 and BnCYP27-4 were the most abundant CYP proteins in leaf extract, only BnCYP18-5 was the most abundant one in phloem sap with a considerably higher relative abundance observed than for the other identified BnCYP proteins from this compartment. Interestingly, the most abundant CYP protein in both compartments is BnCYP18-5. The finding of so many CYPs in the phloem transport system suggests an essential function of this protein family in this compartment, probably in protein transport or long-distance signalling.

## Conclusions

Cyclophilins are ubiquitous proteins that constitute a multigene family in higher organisms. Exceptionally high numbers of CYPs have been found in plants, underlining their essential importance in many essential physiological processes. However, the physiological roles of most CYPs in plants are not well understood.

The present study applied bioinformatic tools for a genome-wide identification of CYPs in the important oil crop *B. napus*. Sequence similarity searches with known Arabidopsis CYPs, sequence alignments and CLD prediction identified a surprisingly high number of 94 CYP-coding genes. Therefore, *B. napus* contains the highest number of CYPs known so far. As in other plants, CYPs are predicted to be localized in all compartments, most of them being probably cytosolic. Most BnCYPs are single-domain proteins.

Transcriptome analysis confirmed the expression of 77 distinct CYPs in the *B. napus* cultivar 'Drakkar' from leaf material under normal growth conditions. The occurrence of 26 BnCYP proteins was confirmed by LC-MS/MS analysis. It is likely that additional CYPs are expressed on transcript and protein levels in different cell types, plant parts, or under different environmental conditions. Interestingly, 12 BnCYPs were exclusively found in phloem samples and not in leaf extract supporting a fundamental and specific role in this specialized compartment.

Future studies must now focus on the functional characterization of the high number of CYPs in *B. napus* in order to better understand the diverse roles of CYPs in oilseed rape and in plant biology in more general. In this regard, elucidating the role(s) of the phloem CYPs in protein refolding and long-distance signalling will be of special interest.

## Materials and Methods

**Sequence analysis.** *Brassica napus* CYPs were identified by BLASTp of the *Arabidopsis thaliana* CYP18-1 (At1g01940) and CYP19-1 (At2g16600) against the *B. napus* genome sequence database<sup>51</sup> deposited at NCBI (<http://www.ncbi.nlm.nih.gov/>)<sup>69</sup>. BLASTp searches with the remaining AtCYPs revealed the same BnCYPs as already identified. Amino acid and cDNA sequences were obtained by the European Nucleotide Archive (<http://www.ebi.ac.uk/ena/data/view/PRJEB5043>). To identify *A. thaliana* homologs for the individual BnCYPs, the amino acid sequences of BnCYPs were used as queries for a BLAST search on UniProtKB (<http://www.uniprot.org/>)<sup>70</sup>.

All identified proteins were analyzed for the presence of a CLD and potential additional domains with InterPro (<http://www.ebi.ac.uk/interpro/>)<sup>71</sup> and drawn by CoreDRAW. The theoretical isoelectric point was determined by the ProteinProspector Tool MS-Digest (<http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msdigest>), subcellular localization was predicted by LocTree3 (<https://roslab.org/services/loctree3/>)<sup>72</sup> and nuclear localization signals (NLS) by using the NLS mapper ([http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS\\_Mapper\\_form.cgi](http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi))<sup>73</sup>.

Chromosome mapping of the CYPs was performed using the *Brassica napus* Genome Browser (<http://www.genoscope.cns.fr/brassicapnapus/cgi-bin/gbrowse/colza/>) available by Genoscope - Centre National de Séquençage and redrawn by CoreDRAW.

**Protein sequence alignment and phylogenetic analysis.** Sequences were aligned using ClustalOmega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>)<sup>74</sup> and displayed with Jalview 2.9.0b2<sup>75</sup>. The secondary structure annotation is based on the structure of the crystallized human CYP A (4N1M.pdb).

The phylogenetic tree was calculated with ClustalW2 Phylogeny ([http://www.ebi.ac.uk/Tools/phylogeny/clustalw2\\_phylogeny/](http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/)) by using the multiple alignment from ClustalOmega, and subsequently processed with iTOL (<http://itol.embl.de/>)<sup>76,77</sup>.

**Plant material and growth conditions.** *Brassica napus* cultivar 'Drakkar' plants were grown in 19 cm pots on soil (LAT-Terra Standard P, Industrie-Erdenwerk Archut, Germany) in a glasshouse under controlled conditions with 70% humidity and a 16 h/8 h light/dark (day/night) and 22 °C/18 °C (day/night) cycle. Plants were watered once per day and fertilized with 2 g/l Osmocote Exact Standard High K (Scotts, the Netherlands).

**Expression profile of cyclophilins.** Transcriptome data were generated by GAMAVIR, a tri-national research activity aiming at characterizing plant:virus interactions in rapeseed (ANR-13-KBBE-0005). PolyA RNA isolated from leaf disks of 6 weeks-old *Brassica napus* cultivar 'Drakkar' plants (n = 3) was sequenced using Illumina technology. Clean RNAseq reads were aligned against the re-sequenced *Brassica napus* cultivar 'Drakkar' genome using Tophat2 v2.0.13<sup>78</sup> and counted with Samtools v1.1<sup>79</sup>.

**Protein extraction for proteomics.** Phloem sap was collected as described previously by Giavalisco *et al.*<sup>2</sup> at the inflorescence of oilseed rape plants before flowering. Phloem sap samples were collected on ice four times

( $n = 4$ ), frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until further processing. Proteins were precipitated in 4 volumes of 90% (v/v) acetone, 10% (v/v) methanol, 10 mM DTT over night at  $-20^{\circ}\text{C}$ . The precipitates were pelleted at  $14,000 \times g$  at  $4^{\circ}\text{C}$  for 15 min, washed twice with 100% acetone and air-dried.

Leaf material was harvested from three different plants ( $n = 3$ ) and for each sample 100 mg material was grinded in liquid nitrogen. Proteins were extracted with 800  $\mu\text{l}$  extraction buffer (50 mM MOPS/pH 7.5, 5% glycerol, 0.55% PVPP, 0.5% Nonidet P-40, 5 mM L-ascorbic acid, 5 mM DTT, 1x protease inhibitor (cComplete Protease Inhibitor Cocktail, Roche), 1x phosphatase inhibitor (PhosStop, Phosphatase Inhibitor Cocktail, Roche). Centrifuging at  $14,000 \times g$  at  $4^{\circ}\text{C}$  for 15 min allowed the separation of soluble proteins from insoluble material. Proteins were precipitated as described above by acetone/methanol/DTT.

**Analysis of phloem sap and leaf proteins by LC-MS/MS.** For the analysis of phloem proteins the extracted and precipitated phloem sap proteins were resuspended in extraction buffer (6 M urea, 2 M thiourea, 15 mM DTT, 2% CHAPS). Here, each protein pellet was dissolved in exactly the same volume deployed for the initial phloem sap precipitation. Once the proteins were in solution, samples were sonicated for 10 min in a sonication bath, followed by 30 min incubation on an orbital shaker (100 rpm) at room temperature. Solubilised proteins were centrifuged at  $10,000 \times g$  for 5 min and the protein concentration was determined from the collected supernatant.

For the analysis of leaf tissue proteins, the precipitated proteins were resuspended in sufficient protein extraction buffer to provide a final protein concentration of  $2 \mu\text{g}/\mu\text{l}$ .

25  $\mu\text{g}$  of phloem sap and 50  $\mu\text{g}$  of leaf tissue protein extract were then digested in solution using a Trypsin/Lys-C mixture (Mass Spec Grade, Promega, Madison, WI, USA) according to the instruction manual. After the digestion, the samples were desalted using C18-stage tips as described in Rappsilber *et al.*<sup>80</sup>.

After the elution of the digested and desalted peptides from C18-stage tips, the samples were concentrated to near dryness in a SpeedVac and the peptide mixtures were reconstituted in 30  $\mu\text{l}$  resuspension buffer (5% acetonitrile, 0.1% formic acid in water). 5  $\mu\text{l}$  of this peptide mix was injected onto an Acclaim PepMap RSLC HPLC column (75  $\mu\text{m} \times 15 \text{ cm}$ , Thermo Scientific) connected to the EASY-nLC 1000 system (Thermo Scientific). The eluting peptides were then analyzed on a Q Exactive Plus (Thermo Scientific, Bremen, Germany) high-resolution mass spectrometer.

The peptides were separated using a binary buffer system of 0.1% formic acid in water (Buffer A) and 60% acetonitrile containing 0.1% formic (Buffer B). The flow rate was adjusted to 300 nl/min. Peptides were eluted with on a linear gradient of 0–40% buffer B for 50 min followed by a linear gradient between 40–80% buffer B for additional 30 min. Peptides were analyzed in the mass spectrometer using one full scan (300–1600 m/z,  $R = 70,000$  at 200 m/z), followed by up to fifteen data-dependent MS/MS scans (Top 15 approach) with higher-energy collisional dissociation (HCD) at a resolution of 17,500 at 200 m/z. Dynamic exclusion was set to 15 s.

Raw data were processed using the Progenesis QI for proteomics (Progenesis QI for Proteomics Version 3.0, Nonlinear Dynamics, Newcastle, UK) software and the protein sequences of all identified CYPs from *Brassica napus*. Protein identifications were filtered with a false discovery rate better than 1%, at least two peptides, one unique peptide and a score of 50.

## References

- Buhtz, A., Pieritz, J., Springer, F. & Kehr, J. Phloem small RNAs, nutrient stress responses, and systemic mobility. *BMC Plant Biology* **10**, 64, doi:10.1186/1471-2229-10-64 (2010).
- Giavalisco, P., Kapitzka, K., Kolasa, A., Buhtz, A. & Kehr, J. Towards the proteome of *Brassica napus* phloem sap. *Proteomics* **6**, 896–909, doi:10.1002/pmic.200500155 (2006).
- Pant, B. D., Buhtz, A., Kehr, J. & Scheible, W. R. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *The Plant Journal* **53**, 731–738, doi:10.1111/tpj.2008.53.issue-5 (2008).
- Kehr, J., Buhtz, A. & Giavalisco, P. Analysis of xylem sap proteins from *Brassica napus*. *BMC Plant Biology* **5**, 11, doi:10.1186/1471-2229-1185-1111 (2005).
- Ostendorp, A., Pahlow, S., Deke, J., Thieß, M. & Kehr, J. Protocol: optimisation of a grafting protocol for oilseed rape (*Brassica napus*) for studying long-distance signalling. *Plant Methods* **12**, 1–8, doi:10.1186/s13007-016-0122-x (2016).
- Kagale, S. *et al.* Polyploid evolution of the Brassicaceae during the Cenozoic era. *Plant Cell* **26**, 2777–2791, doi:10.1105/tpc.114.126391 (2014).
- Schmutzer, T. *et al.* Species-wide genome sequence and nucleotide polymorphisms from the model allopolyploid plant *Brassica napus*. *Scientific data* **2**, 150072, doi:10.1038/sdata.2015.72 (2015).
- Íñiguez-Luy, F. L. F., M. L. L. In *Genetics and Genomics of the Brassicaceae* (ed I.; Schmidt Bancroft, R.) 291–322 (New York, NY: Springer, 2011).
- Friedt, W. & Snowdon, R. In *Oil Crops Vol. 4 Handbook of Plant Breeding* (eds Johann Vollmann & Istvan Rajcan) Ch. 4, 91–126 (Springer New York, 2010).
- Harding, M. W., Galat, A., Uehling, D. E. & Schreiber, S. L. A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. *Nature* **341**, 758–760, doi:10.1038/341758a0 (1989).
- Bierer, B. E. *et al.* Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin. *Proc Natl Acad Sci USA* **87**, 9231–9235, doi:10.1073/pnas.87.23.9231 (1990).
- Handschumacher, R. E., Harding, M. W., Rice, J., Drugge, R. J. & Speicher, D. W. Cyclophilin: a specific cytosolic binding protein for cyclosporin A. *Science* **226**, 544–547, doi:10.1126/science.6238408 (1984).
- Harding, M. W., Handschumacher, R. E. & Speicher, D. W. Isolation and amino acid sequence of cyclophilin. *J Biol Chem* **261**, 8547–8555 (1986).
- Kieffhaber, T., Quaas, R., Hahn, U. & Schmid, F. X. Folding of ribonuclease T1. 2. Kinetic models for the folding and unfolding reactions. *Biochemistry* **29**, 3061–3070, doi:10.1021/bi00464a024 (1990).
- Rahfeld, J. U. *et al.* Confirmation of the existence of a third family among peptidyl-prolyl cis/trans isomerases. Amino acid sequence and recombinant production of parvulin. *FEBS Lett* **352**, 180–184, doi:10.1016/0014-5793(94)00932-5 (1994).
- Galat, A. Peptidylprolyl cis/trans isomerases (immunophilins): biological diversity—targets—functions. *Current topics in medicinal chemistry* **3**, 1315–1347, doi:10.2174/1568026033451862 (2003).
- Liu, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807–815, doi:10.1016/0092-8674(91)90124-H (1991).

18. Barik, S. Immunophilins: for the love of proteins. *Cell Mol Life Sci* **63**, 2889–2900, doi:10.1007/s00018-006-6215-3 (2006).
19. Gasser, C. S., Gunning, D. A., Budelier, K. A. & Brown, S. M. Structure and expression of cytosolic cyclophilin/peptidyl-prolyl cis-trans isomerase of higher plants and production of active tomato cyclophilin in *Escherichia coli*. *Proc Natl Acad Sci USA* **87**, 9519–9523, doi:10.1073/pnas.87.24.9519 (1990).
20. Galat, A. Variations of sequences and amino acid compositions of proteins that sustain their biological functions: An analysis of the cyclophilin family of proteins. *Arch Biochem Biophys* **371**, 149–162, doi:10.1006/abbi.1999.1434 (1999).
21. Klappa, P., Freedman, R. B. & Zimmermann, R. Protein disulphide isomerase and a luminal cyclophilin-type peptidyl prolyl cis-trans isomerase are in transient contact with secretory proteins during late stages of translocation. *Eur J Biochem* **232**, 755–764, doi:10.1111/ejb.1995.232.issue-3 (1995).
22. Kern, G., Kern, D., Schmid, F. X. & Fischer, G. A kinetic analysis of the folding of human carbonic anhydrase II and its catalysis by cyclophilin. *J Biol Chem* **270**, 740–745, doi:10.1074/jbc.270.2.740 (1995).
23. Zander, K. *et al.* Cyclophilin A interacts with HIV-1 Vpr and is required for its functional expression. *J Biol Chem* **278**, 43202–43213, doi:10.1074/jbc.M305414200 (2003).
24. Baker, E. K., Colley, N. J. & Zuker, C. S. The cyclophilin homolog NinaA functions as a chaperone, forming a stable complex in vivo with its protein target rhodopsin. *EMBO J* **13**, 4886–4895 (1994).
25. Allain, F., Denys, A. & Spik, G. Characterization of surface binding sites for cyclophilin B on a human tumor T-cell line. *J Biol Chem* **269**, 16537–16540 (1994).
26. Schiene-Fischer, C. & Yu, C. Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl cis/trans isomerases. *FEBS Lett* **495**, 1–6, doi:10.1016/S0014-5793(01)02326-2 (2001).
27. Brazin, K. N., Mallis, R. J., Fulton, D. B. & Andreotti, A. H. Regulation of the tyrosine kinase Itk by the peptidyl-prolyl isomerase cyclophilin A. *Proc Natl Acad Sci USA* **99**, 1899–1904, doi:10.1073/pnas.042529199 (2002).
28. Bukrinsky, M. I. Cyclophilins: unexpected messengers in intercellular communications. *Trends Immunol* **23**, 323–325, doi:10.1016/S1471-4906(02)02237-8 (2002).
29. Jing, H. *et al.* Peptidyl-prolyl isomerization targets rice Aux/IAAs for proteasomal degradation during auxin signalling. *Nat Commun* **6**, 7395, doi:10.1038/ncomms8395 (2015).
30. Pogorelko, G. V. *et al.* Characterization of three Arabidopsis thaliana immunophilin genes involved in the plant defense response against *Pseudomonas syringae*. *Gene* **538**, 12–22, doi:10.1016/j.gene.2014.01.029 (2014).
31. Lin, D. T. & Lechleiter, J. D. Mitochondrial targeted cyclophilin D protects cells from cell death by peptidyl prolyl isomerization. *J Biol Chem* **277**, 31134–31141, doi:10.1074/jbc.M112035200 (2002).
32. Krzywicka, A. *et al.* KIN241: a gene involved in cell morphogenesis in *Paramecium tetraurelia* reveals a novel protein family of cyclophilin-RNA interacting proteins (CRIPs) conserved from fission yeast to man. *Mol Microbiol* **42**, 257–267, doi:10.1046/j.1365-2958.2001.02634.x (2001).
33. Anderson, M. *et al.* A new family of cyclophilins with an RNA recognition motif that interact with members of the trx/MLL protein family in *Drosophila* and human cells. *Dev Genes Evol* **212**, 107–113, doi:10.1007/s00427-002-0213-8 (2002).
34. Dubourg, B. *et al.* The human nuclear SRcyp is a cell cycle-regulated cyclophilin. *J Biol Chem* **279**, 22322–22330, doi:10.1074/jbc.M400736200 (2004).
35. Li, H. *et al.* A WD40 domain cyclophilin interacts with histone H3 and functions in gene repression and organogenesis in Arabidopsis. *Plant Cell* **19**, 2403–2416, doi:10.1105/tpc.107.053579 (2007).
36. Bannikova, O. *et al.* Identification of RNA targets for the nuclear multidomain cyclophilin atCyp59 and their effect on PPIase activity. *Nucleic Acids Res* **41**, 1783–1796, doi:10.1093/nar/gks1252 (2013).
37. Ahn, J. C. *et al.* Classification of rice (*Oryza sativa* L. Japonica nipponbare) immunophilins (FKBPs, CYPs) and expression patterns under water stress. *BMC Plant Biol* **10**, 253, doi:10.1186/1471-2229-10-253 (2010).
38. Dominguez-Solis, J. R. *et al.* A cyclophilin links redox and light signals to cysteine biosynthesis and stress responses in chloroplasts. *Proc Natl Acad Sci USA* **105**, 16386–16391, doi:10.1073/pnas.0808204105 (2008).
39. Ruan, S. L. *et al.* Proteomic identification of OsCYP2, a rice cyclophilin that confers salt tolerance in rice (*Oryza sativa* L.) seedlings when overexpressed. *BMC Plant Biol* **11**, 34, doi:10.1186/1471-2229-11-34 (2011).
40. Park, S. W. *et al.* Cyclophilin 20-3 relays a 12-oxo-phytyldienoic acid signal during stress responsive regulation of cellular redox homeostasis. *Proc Natl Acad Sci USA* **110**, 9559–9564, doi:10.1073/pnas.1218872110 (2013).
41. Gottschalk, M. *et al.* Ricinus communis cyclophilin: functional characterisation of a sieve tube protein involved in protein folding. *Planta* **228**, 687–700, doi:10.1007/s00425-008-0771-8 (2008).
42. Esau, K. *The phloem*. Vol. 5 (Gebr. Borntraeger, 1969).
43. Kehr, J. Long-distance transport of macromolecules through the phloem. *F1000 Biol Rep* **1**, 31, doi:10.3410/B1-31 (2009).
44. Galat, A. Function-dependent clustering of orthologues and paralogues of cyclophilins. *Proteins* **56**, 808–820, doi:10.1002/prot.20156 (2004).
45. Pemberton, T. J. & Kay, J. E. The cyclophilin repertoire of the fission yeast *Schizosaccharomyces pombe*. *Yeast* **22**, 927–945, doi:10.1002/yea.1288 (2005).
46. Singh, K., Zouhar, M., Mazakova, J. & Rysanek, P. Genome wide identification of the immunophilin gene family in *Leptosphaeria maculans*: a causal agent of Blackleg disease in Oilseed Rape (*Brassica napus*). *OMICS* **18**, 645–657, doi:10.1089/omi.2014.0081 (2014).
47. Romano, P. G., Horton, P. & Gray, J. E. The Arabidopsis cyclophilin gene family. *Plant Physiol* **134**, 1268–1282, doi:10.1104/pp.103.022160 (2004).
48. He, Z., Li, L. & Luan, S. Immunophilins and parvulins. Superfamily of peptidyl prolyl isomerases in Arabidopsis. *Plant Physiol* **134**, 1248–1267, doi:10.1104/pp.103.031005 (2004).
49. Vasudevan, D. *et al.* Plant immunophilins: a review of their structure-function relationship. *Biochim Biophys Acta* **1850**, 2145–58, doi:10.1016/j.bbagen.2014.12.017 (2014).
50. Mainali, H. R., Chapman, P. & Dhaubhadel, S. Genome-wide analysis of Cyclophilin gene family in soybean (*Glycine max*). *BMC Plant Biol* **14**, 282, doi:10.1186/s12870-014-0282-7 (2014).
51. Chalhoub, B. *et al.* Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* **345**, 950–953, doi:10.1126/science.1253435 (2014).
52. Ke, H. M., Zydowsky, L. D., Liu, J. & Walsh, C. T. Crystal structure of recombinant human T-cell cyclophilin A at 2.5 Å resolution. *Proc Natl Acad Sci USA* **88**, 9483–9487, doi:10.1073/pnas.88.21.9483 (1991).
53. Zydowsky, L. D. *et al.* Active site mutants of human cyclophilin A separate peptidyl-prolyl isomerase activity from cyclosporin A binding and calcineurin inhibition. *Protein Sci* **1**, 1092–1099, doi:10.1002/pro.5560010903 (1992).
54. Liu, J., Chen, C. M. & Walsh, C. T. Human and *Escherichia coli* cyclophilins: sensitivity to inhibition by the immunosuppressant cyclosporin A correlates with a specific tryptophan residue. *Biochemistry* **30**, 2306–2310, doi:10.1021/bi00223a003 (1991).
55. Lee, S. S. *et al.* Rice cyclophilin OsCYP18-2 is translocated to the nucleus by an interaction with SKIP and enhances drought tolerance in rice and Arabidopsis. *Plant Cell Environ* **38**, 2071–87, doi:10.1111/pce.12531 (2015).
56. Iki, T., Yoshikawa, M., Meshi, T. & Ishikawa, M. Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. *EMBO J* **31**, 267–278, doi:10.1038/emboj.2011.395 (2012).
57. Earley, K. W. & Poethig, R. S. Binding of the cyclophilin 40 ortholog SQUINT to Hsp90 protein is required for SQUINT function in Arabidopsis. *J Biol Chem* **286**, 38184–38189, doi:10.1074/jbc.M111.290130 (2011).

58. Ratajczak, T. & Carrello, A. Cyclophilin 40 (CyP-40), mapping of its hsp90 binding domain and evidence that FKBP52 competes with CyP-40 for hsp90 binding. *J Biol Chem* **271**, 2961–2965, doi:10.1074/jbc.271.6.2961 (1996).
59. Nelson, C. J., Santos-Rosa, H. & Kouzarides, T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* **126**, 905–916, doi:10.1016/j.cell.2006.07.026 (2006).
60. Li, H. & Luan, S. The cyclophilin AtCYP71 interacts with CAF-1 and LHP1 and functions in multiple chromatin remodeling processes. *Mol Plant* **4**, 748–758, doi:10.1093/mp/ssr036 (2011).
61. Vasudevan, D., Fu, A., Luan, S. & Swaminathan, K. Crystal structure of Arabidopsis cyclophilin38 reveals a previously uncharacterized immunophilin fold and a possible autoinhibitory mechanism. *Plant Cell* **24**, 2666–2674, doi:10.1105/tpc.111.093781 (2012).
62. Shapiguzov, A., Edvardsson, A. & Vener, A. V. Profound redox sensitivity of peptidyl-prolyl isomerase activity in Arabidopsis thylakoid lumen. *FEBS Lett* **580**, 3671–3676, doi:10.1016/j.febslet.2006.05.054 (2006).
63. Edvardsson, A., Eshaghi, S., Vener, A. V. & Andersson, B. The major peptidyl-prolyl isomerase activity in thylakoid lumen of plant chloroplasts belongs to a novel cyclophilin TLP20. *FEBS Lett* **542**, 137–141, doi:10.1016/S0014-5793(03)00366-1 (2003).
64. Fu, A. *et al.* A chloroplast cyclophilin functions in the assembly and maintenance of photosystem II in Arabidopsis thaliana. *Proc Natl Acad Sci USA* **104**, 15947–15952, doi:10.1073/pnas.0707851104 (2007).
65. Sirpio, S. *et al.* AtCYP38 ensures early biogenesis, correct assembly and sustenance of photosystem II. *Plant J* **55**, 639–651, doi:10.1111/j.1365-313X.2008.03532.x (2008).
66. Weighardt, F. *et al.* A novel hnRNP protein (HAP/SAF-B) enters a subset of hnRNP complexes and relocates in nuclear granules in response to heat shock. *J Cell Sci* **112** (Pt 10), 1465–1476 (1999).
67. Gullerova, M., Barta, A. & Lorkovic, Z. J. AtCyp59 is a multidomain cyclophilin from Arabidopsis thaliana that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. *RNA* **12**, 631–643, doi:10.1261/rna.2226106 (2006).
68. Cheng, F., Wu, J. & Wang, X. Genome triplication drove the diversification of Brassica plants. *Horticulture research* **1**, 14024, doi:10.1038/hortres.2014.24 (2014).
69. Coordinators, N. R. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**, D7–19, doi:10.1093/nar/gkv1290 (2016).
70. UniProt, C. UniProt: a hub for protein information. *Nucleic Acids Res* **43**, D204–212, doi:10.1093/nar/gku989 (2015).
71. Mitchell, A. *et al.* The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res* **43**, D213–221, doi:10.1093/nar/gku1243 (2015).
72. Goldberg, T. *et al.* LocTree3 prediction of localization. *Nucleic Acids Res* **42**, W350–355, doi:10.1093/nar/gku396 (2014).
73. Kosugi, S., Hasebe, M., Tomita, M. & Yanagawa, H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci USA* **106**, 10171–10176, doi:10.1073/pnas.0900604106 (2009).
74. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539–539, doi:10.1038/msb.2011.75 (2011).
75. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191, doi:10.1093/bioinformatics/btp033 (2009).
76. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128, doi:10.1093/bioinformatics/btl529 (2007).
77. Letunic, I. & Bork, P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39**, W475–478, doi:10.1093/nar/gkr201 (2011).
78. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).
79. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079, doi:10.1093/bioinformatics/btp352 (2009).
80. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896–1906, doi:10.1038/nprot.2007.261 (2007).
81. Sakuma, Y. *et al.* Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proc Natl Acad Sci USA* **103**, 18822–18827, doi:10.1073/pnas.0605639103 (2006).
82. Trupkin, S. A., Mora-Garcia, S. & Casal, J. J. The cyclophilin ROC1 links phytochrome and cryptochrome to brassinosteroid sensitivity. *Plant J* **71**, 712–723, doi:10.1111/j.1365-313X.2012.05013.x (2012).
83. Li, M. *et al.* Proline isomerization of the immune receptor-interacting protein RIN4 by a cyclophilin inhibits effector-triggered immunity in Arabidopsis. *Cell Host Microbe* **16**, 473–483, doi:10.1016/j.chom.2014.09.007 (2014).
84. Kovalev, N. & Nagy, P. D. Cyclophilin A binds to the viral RNA and replication proteins, resulting in inhibition of tombusviral replicase assembly. *J Virol* **87**, 13330–13342, doi:10.1128/JVI.02101-13 (2013).
85. Deng, W. *et al.* Agrobacterium VirD2 protein interacts with plant host cyclophilins. *Proc Natl Acad Sci USA* **95**, 7040–7045, doi:10.1073/pnas.95.12.7040 (1998).
86. Ma, X., Song, L., Yang, Y. & Liu, D. A gain-of-function mutation in the ROC1 gene alters plant architecture in Arabidopsis. *New Phytol* **197**, 751–762, doi:10.1111/nph.12056 (2013).
87. Coaker, G., Fallick, A. & Staskawicz, B. Activation of a phytopathogenic bacterial effector protein by a eukaryotic cyclophilin. *Science* **308**, 548–550, doi:10.1126/science.1108633 (2005).
88. Coaker, G., Zhu, G., Ding, Z., Van Doren, S. R. & Staskawicz, B. Eukaryotic cyclophilin as a molecular switch for effector activation. *Mol Microbiol* **61**, 1485–1496, doi:10.1111/j.1365-2958.2006.05335.x (2006).
89. Aumuller, T., Jahreis, G., Fischer, G. & Schiene-Fischer, C. Role of prolyl cis/trans isomers in cyclophilin-assisted Pseudomonas syringae AvrRpt2 protease activation. *Biochemistry* **49**, 1042–1052, doi:10.1021/bi901813e (2010).
90. Stangeland, B. *et al.* Molecular analysis of Arabidopsis endosperm and embryo promoter trap lines: reporter-gene expression can result from T-DNA insertions in antisense orientation, in introns and in intergenic regions, in addition to sense insertion at the 5' end of genes. *J Exp Bot* **56**, 2495–2505, doi:10.1093/jxb/eri242 (2005).
91. Kaur, G. *et al.* Characterization of Peptidyl-Prolyl Cis-Trans Isomerase- and Calmodulin-Binding Activity of a Cytosolic Arabidopsis thaliana Cyclophilin AtCyp19-3. *PLoS one* **10**, e0136692, doi:10.1371/journal.pone.0136692 (2015).
92. Saito, T. *et al.* Expression of a gene for cyclophilin which contains an amino-terminal endoplasmic reticulum-targeting signal. *Plant Cell Physiol* **40**, 77–87, doi:10.1093/oxfordjournals.pcp.a029477 (1999).
93. Grebe, M. *et al.* A conserved domain of the Arabidopsis GNOM protein mediates subunit interaction and cyclophilin 5 binding. *Plant Cell* **12**, 343–356, doi:10.1105/tpc.12.3.343 (2000).
94. Kamauchi, S., Nakatani, H., Nakano, C. & Urade, R. Gene expression in response to endoplasmic reticulum stress in Arabidopsis thaliana. *FEBS J* **272**, 3461–3476, doi:10.1111/j.1742-4658.2005.04770.x (2005).
95. Jackson, K. & Soll, D. Mutations in a new Arabidopsis cyclophilin disrupt its interaction with protein phosphatase 2A. *Mol Gen Genet* **262**, 830–838, doi:10.1007/s004380051147 (1999).
96. Schubert, M. *et al.* Proteome map of the chloroplast lumen of Arabidopsis thaliana. *J Biol Chem* **277**, 8354–8365, doi:10.1074/jbc.M108575200 (2002).
97. Ingelsson, B., Shapiguzov, A., Kieselbach, T. & Vener, A. V. Peptidyl-prolyl isomerase activity in chloroplast thylakoid lumen is a dispensable function of immunophilins in Arabidopsis thaliana. *Plant Cell Physiol* **50**, 1801–1814, doi:10.1093/pcp/pcp122 (2009).



98. Zhang, Y. *et al.* The cyclophilin CYP20-2 modulates the conformation of BRASSINAZOLE-RESISTANT1, which binds the promoter of FLOWERING LOCUS D to regulate flowering in Arabidopsis. *Plant Cell* **25**, 2504–2521, doi:10.1105/tpc.113.110296 (2013).
99. Sirpio, S., Holmstrom, M., Battchikova, N. & Aro, E. M. AtCYP20-2 is an auxiliary protein of the chloroplast NAD(P)H dehydrogenase complex. *FEBS Lett* **583**, 2355–2358, doi:10.1016/j.febslet.2009.06.031 (2009).
100. Edvardsson, A., Shapiguzov, A., Petersson, U. A., Schroder, W. P. & Vener, A. V. Immunophilin AtFKBP13 sustains all peptidyl-prolyl isomerase activity in the thylakoid lumen from Arabidopsis thaliana deficient in AtCYP20-2. *Biochemistry* **46**, 9432–9442, doi:10.1021/bi700426q (2007).
101. Lippuner, V. *et al.* Cloning and characterization of chloroplast and cytosolic forms of cyclophilin from Arabidopsis thaliana. *J Biol Chem* **269**, 7863–7868 (1994).
102. Motohashi, K., Koyama, E., Nakanishi, Y., Ueoka-Nakanishi, H. & Hisabori, T. Chloroplast cyclophilin is a target protein of thioredoxin. *J Biol Chem* **278**, 31848–31852, doi:10.1074/jbc.M304258200 (2003).
103. Laxa, M., Konig, J., Dietz, K. J. & Kandlbinder, A. Role of the cysteine residues in Arabidopsis thaliana cyclophilin CYP20-3 in peptidyl-prolyl cis-trans isomerase and redox-related functions. *Biochem J* **401**, 287–297, doi:10.1042/BJ20061092 (2007).
104. Liebthal, M. *et al.* Redox-Dependent Conformational Dynamics of Decameric 2-Cysteine Peroxiredoxin and its Interaction with Cyclophilin 20-3. *Plant and Cell Physiology* **57**, 1415–1425, doi:10.1093/pcp/pcw031 (2016).
105. Kim, J. H., Nguyen, N. H., Nguyen, N. T., Hong, S. W. & Lee, H. Loss of all three calreticulins, CRT1, CRT2 and CRT3, causes enhanced sensitivity to water stress in Arabidopsis. *Plant Cell Rep* **32**, 1843–1853, doi:10.1007/s00299-013-1497-z (2013).
106. Tomasic Paic, A. & Fulgosi, H. Chloroplast immunophilins. *Protoplasma* **253**, 249–258, doi:10.1007/s00709-015-0828-z (2015).
107. Smith, M. R. *et al.* Cyclophilin 40 is required for microRNA activity in Arabidopsis. *Proc Natl Acad Sci USA* **106**, 5424–5429, doi:10.1073/pnas.0812729106 (2009).

## Acknowledgements

We are grateful to S. Graindorge and D. Pflieger at the IBMP BioImage and Bioinformatics platform for providing access to the bioinformatic resources and the analysis of RNAseq data. We would like to acknowledge the financial contribution to the research activities by a Career Integration Grant (CIG; PCIG14-GA-2013-63 0734) by the European Commission within the 7th framework program, the grant LFF-GK06 'DELIGRAH' (Landesforschungsförderung Hamburg), and a DFG grant (DFG KE 856\_6-1) awarded to J.K., and funding by the Agence National de la Recherche (ANR-13-KBBE-0005-01) awarded to M.H.

## Author Contributions

P.H. and M.T. performed the original database searches and contributed to the design of the project. P.H. performed bioinformatic predictions, phloem sap sampling and protein extraction. K.A. performed RNA extraction, K.A. and M.H. contributed to RNA-Seq data analysis. K.B. and P.G. performed LC-MS/MS. P.H. and P.G. analyzed the protein data. J.K. and M.H. participated in data analysis and discussion. P.H. and M.T. drafted the initial manuscript, J.K. conceived of and supervised the study and complemented the writing. All authors reviewed the manuscript and approved the final article.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-01596-5

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017