

# Impact of conventional breeding on the whole seed proteome of *Brassica napus* L. using quantitative differential 2D electrophoresis and shotgun proteomics

Devouge Vanessa<sup>1</sup>, Rogniaux Hélène<sup>1</sup>, Tessier Dominique<sup>1</sup>, Gilbert Deshayes<sup>1</sup>, Guéguen Jacques<sup>1</sup>, Cécile Baron<sup>2</sup>, Nési Nathalie<sup>2</sup>, Larré Colette<sup>1</sup>

<sup>1</sup> INRA, BIA, Rue de la Géraudière, BP 71627, 44316 Nantes, France

<sup>2</sup> INRA Rennes, APBV, BP 35327, 35653 Le Rheu, France Email: larre@nantes.inra.fr

## Abstract

In this study, four near-isogenic *B.napus* varieties with absence of erucic acid and decreasing amounts glucosinolates were used to characterise the proteins affected during the breeding process. Two strategies were investigated, 1/ by quantitative differential 2D electrophoresis and 2/ by shotgun proteomic approach involving isobaric tagging of peptides.

1) Using 2D electrophoresis analysis, 72 spots were found to be differentially expressed between the studied lines. Forty were identified by mass spectrometry that are involved mainly into carbohydrate pathway or detoxification/ defence. A validation of the differential expression of two of these proteins was performed using quantitative PCR.

2) Despite the high performance of 2D electrophoresis, it was difficult to quantify the storage proteins. Cruciferins were scattered into at least 20 distinct spots and napins were hardly detected due to their low molecular mass and the basicity of some isoforms. In this context, we used an alternative shotgun proteomic approach which is based on the enzymatic hydrolysis of the entire proteome, followed by a multi-dimensional chromatographic separation of the resulting peptides. In order to compare the expression level of proteins between the four lines, peptides were labeled with iTRAQs reagents prior to tandem mass spectrometry. Results show that the ratio between cruciferins and napins in mature seeds correlates to the amount of glucosinolates.

**Key words:** Oilseed rape, *Brassica napus*, 2D electrophoresis, breeding, napin, cruciferin, shotgun proteomic, iTRAQ

## Introduction

Rapeseed was subjected to intensive selective breeding to reach what is called canola-quality rapeseed (double-low “00”, “zero erucic acid, zero glucosinolates”) with the reduction in two toxic compounds : the erucic acid in the edible oil was eliminated, and the glucosinolate level was decreased in the meal for feedstock uses. *Brassica napus* seeds contain a series of glucosinolates derived from various amino acids. These secondary metabolites are involved in defence against predators or parasites through their degradation products (Bone & Rossiter, 1996) which are also toxic in nutrition. After rapeseed industrial transformation, glucosinolates and their degradation products still remain in meal. The improvements made through breeding research programs paid little attention to the seed protein composition. However previous studies showed that the balance in the major storage proteins (napins (2S) and cruciferins (12S)) has been modified into the 00 varieties (Raab et al., 1992; Malabat et al., 2003).

Previous studies comparing varieties by 2D electrophoresis pointed out a large number of protein variations, new isoforms, displacement of protein probably related to different post translational modifications and a very large number of quantitative differences. In the case of *Arabidopsis*, a proteomic comparison of eight ecotypes revealed that they displayed qualitative as well as quantitative differences. They share only 25% of their total spots and 10 % of them were specific for one ecotype (Chevalier et al., 2004). The number of common spots increased considerably when the comparison is achieved between varieties of the same species. Two *B. oleracea* were shown to share around 77 % of their leaf proteome and 78 % of their stem proteome (calculated from Albertin’s data, 2005).

In our study, we decided to explore the impact of the breeding process (reduction of erucic acid and glucosinolate contents) on the seed proteome by using two strategies : 1) quantitative differential 2D electrophoresis and 2) shotgun proteomic approach involving isobaric tagging of peptides. Despite the well known efficiency of 2D electrophoresis, the second strategy was developed in order to get information on the 2S storage protein (napin) which, because of their high basicity added to their small molecular weight, are not detected in 2D gels.

## Material and Methods

**Plant material and growth conditions.** Four winter-type *Brassica napus* varieties, namely Gaspard (++), JetNeuf (0+), Darmor (00) and Darmor-bzh (00) that belong to a near-isogenic family, provided by APBV INRA Research Centre in Rennes, France, were used in this study (Renard et al., 2001). Plants were grown under controlled and reproducible conditions (16-hours photoperiod; 18 to 20°C night/20 to 22°C day temperature). Dry mature seeds were harvested at 60 days after pollination.

**2D electrophoresis.** Dry mature *Brassica napus* seeds were ground in a buffer (8M Urea, 2% Chaps, 2M Thiourea,

18mM DTT, 2% ASB C80). After 2 hours agitation at room temperature, the extracts were centrifuged at 13,000g for 30 min and the protein fraction was recovered in the supernatant.

Isoelectrofocusing was performed on strips pH 3-10 loaded with 200 or 500 µg of proteins. The proteins were separated in a second dimension on 10% and 15% polyacrylamide gels. The gels were stained with Coomassie Brilliant Blue G250 (Sigma-Aldrich) according to Consoli & Damerval, 2001. The images obtained from the four varieties were analysed using Image Master 2D Platinum software (GE Healthcare). The statistical analysis were performed using Statgraphics. Protein spots were picked up manually and in gel digested by trypsin prior to mass spectrometry. The resulting peptide mixture was acidified by the addition of 1 µL of an aqueous solution of formic acid (1%, vol.).

Sample preparation and labeling for shotgun analysis. One hundred micrograms of protein from each sample were reduced, alkylated, digested with trypsin and labeled with the isobaric reagents according to the protocol given in the iTRAQ™ reagent kit (Applied Biosystems). Labeled samples were pooled and subjected to strong cation exchange (PolyLC Polysulfoethyl A column (4.6 mm x 100 mm) fractionation with salt gradient (Agilent Technologies, Inc, Palo Alto, CA, USA). Collected fractions were dried using a vacuum centrifuge. Samples were labeled as follows: 114 tag – Darmor<sub>bzh</sub>; 115 tag – Darmor; 116 tag – Jetneuf; 117 tag – Gaspard

*Protein identification by mass spectrometry.* The digested proteins and the labelled peptides mixture were analysed on a nanoscale capillary liquid chromatography-tandem mass spectrometry (LC-MS/MS) using a Switchos-Ultimate capillary LC system (LC Packings/Dionex, Amsterdam, the Netherlands), coupled with a hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Q-TOF Global, Micromass/Waters, Manchester, UK). Mass data acquisitions were processed with Masslynx software (Micromass/Waters) using the so-called “data dependent acquisition” mode. Protein identification was achieved by searching mass data in the UniProt/Swiss-Prot and UniProt/TrEMBL databanks (10-01-2006), or in the TIGR Gene Indices databank (Brassica napus: release: 29-09-2004).

*Gene expression quantification using real-time quantitative PCR.* Total RNA was extracted with the SV Total RNA Isolation System (Promega, Madison, USA) according to the instructions of the manufacturer. For reverse transcription (RT)-PCR studies, 2.5 µg of DNA-free RNA extract was converted into first-strand cDNA by using the SuperScriptII preamplification system for first-strand synthesis kit (Invitrogen) and oligo(dT)<sub>12-18</sub>. Amplification of cDNA was conducted in optical 96-well plates with an ABI PRISM® 7700 cyclor (Perkin-Elmer Applied Biosystems, Foster City, CA) using SYBR® Green to monitor dsDNA synthesis. Reactions were performed with 12.5 µL 2x SYBR Green Master Mix reagent (Applied Biosystems), 12.5 ng of cDNA and 900 nM of each primers in a final volume of 25 µL, under the following thermal conditions : 50°C for 2 min; 95°C for 10 min; 40 cycles of 95°C for 15 sec and 60°C for 1 min. At the end of the reaction, a dissociation curve (95°C for 15 sec; 60°C for 20 sec; a 20 min-ramp to 95°C; 95°C for 15 sec) was run to verify amplicon specificity. Data were analysed using the SDS 1.7 software (Perkin-Elmer Applied Biosystems). To determine the level of specific transcripts present in a sample, we performed relative quantification through the method of the standard curve (according to ABI7700 user guide). Gene specific primers were designed with Primer Express 1.5 software (Perkin-Elmer Applied Biosystems) using the basic parameters (primer size 18-25 bases; 20-80% GC; primer Tm 58-62°C; amplicon length 50-150 bp; amplicon Tm > 76°C). Sequences are as follows:

Q-BnEF1-1-UP 5'-CGGTGTCATCAAGAGCGTTG-3'; Q-BnEF1-1-RP 5'-GGACTTTGATTCATCTGCGGAT-3';  
 b-glucosidase-UP 5'-CCGGGCTAAAACAGAAATCAT-3'; b-glucosidase-RP  
 5'-TGAGAAGAACTCCAACACAACTCA-3'; Glutathione S transferase-UP 5'-CCCTGGGTATGAGAAGTTTGGT-3';  
 Glutathione S transferase -RP 5'-CCTCTTAGCCCATGCAATCAG-3'

## Results and discussion

### *2D electrophoresis reveals slight differences in protein expression within the near isogenic line*

The seed proteomes of the four varieties were compared using 200 µg protein loading and 15% acrylamide gels. A mean number of spots of 180 was obtained in these conditions. Five very intense spots were present on these gels; using western blot analysis they were identified as cruciferin polypeptides. This identification has been confirmed and refined by mass spectrometry. The cruciferins polypeptides were scattered into about 20 spots as revealed by Western blot analysis. The sum of the volumes of these spots accounted for about 30 % of the total spots volume on the gels obtained for all the varieties. Only two of them identified as cruciferin beta polypeptides were differentially expressed between the varieties.

In order to get deeper into the seed proteome, the varieties were compared using 500 µg of protein loading and 10% acrylamide gels. These conditions permitted to increase the number of spots. A slightly but significantly different number of spots was detected between Gaspard, JetNeuf, Darmor and Darmor-*bzh* with 495 ± 40, 456 ± 29, 562 ± 8 and 604 ± 32 spots, respectively. A multiple range test at 0.01 confidence level revealed a total of 69 differences between these varieties.

Gaspard (++) and Jetneuf (0+) differ by their acid erucic content, and present 55 different spots and most of them were over expressed in Jetneuf. Among these, 32 proteins were identified and are mainly involved in two functional classes : carbohydrate pathway and detoxification/ defence system.

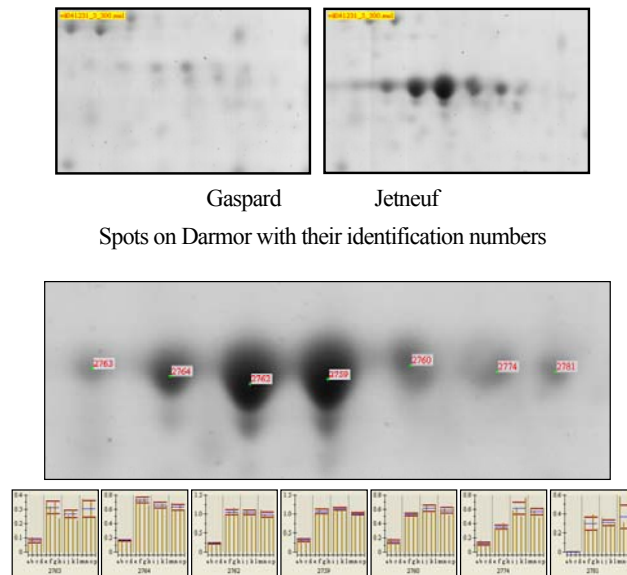


Fig 1 : Zoom on seven spots of TC358 homologue to beta glucosidase Q42618 in a 2D gel electrophoresis.

Expression pattern of each spot across the four varieties : Gaspard, Jetneuf, Darmor, Darmor bzh.

Beta-glucosidase class 1 (TC 358 in the TIGR gene indices) is an example of proteins expressed differentially between the varieties (Figure 1). Beta glucosidase is implicated in the polysaccharide catabolism and such a large variation in its expression between Gaspard and its progeny was not expected.

Therefore we quantified its mRNA expression by RT PCR within the four varieties during the last developmental stages (Figure 2).

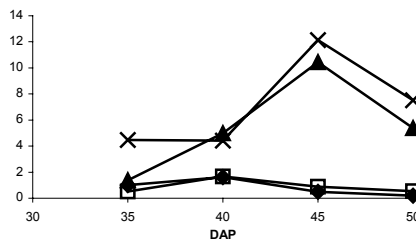


Fig 2 : Relative expression profiles of beta glucosidase (Q42618) measured in Gaspard (rhombus), Jetneuf (squares), Darmor (triangles) and Darmor bzh (crosses)

The expression patterns of beta glucosidase in the mature seeds were quite similar across the seven spots, this protein is clearly under expressed in Gaspard. The comparison with mRNA expression patterns shows that the mRNA level is very low for Gaspard and higher for Darmor and Darmor<sub>bzh</sub>. However the mRNA expression pattern in JetNeuf is comparable to that in Gaspard. It seems that in this case, the transcripts are present but not their products.

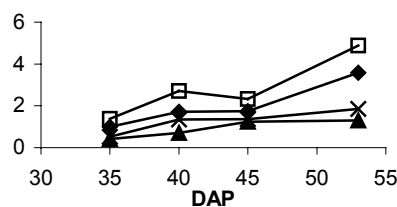


Fig 3 : Expression profiles of glutathione S transferase (Q9ZRW8) measured in Gaspard (rhombus), Jetneuf (squares), Darmor (triangles) and Darmor bzh (crosses)

The differential expression of beta glucosidase within the near isogenic line was confirmed. The relation with the breeding process is quite unclear, although a potential involvement of these enzymes in stress response and plant defence has been recently suggested (Stotz et al., 2000; Matsushima et al., 2003),.

Among the proteins involved in the defence/detoxification system, we focused here on a glutathione S transferase (GST, Q9ZRW8) over expressed in double null varieties compared to their parents. In this case a good correlation was found with the corresponding mRNA patterns (Figure 3). Many functions have been described for GSTs : conjugation between G-SH and xenobiotics, defence against oxidative stress and involvement in plant cell stress signaling. Taking into account these functions

and especially those of detoxification, it cannot be excluded that the less amount of GST in double null lines could be a consequence of the breeding process.

As expected, the use of varieties belonging to a near-isogenic family lead to a reduction of the proteome variability. Lowering the background we found 39 differentially expressed proteins however in most of cases it was difficult to rely their expression and their functions to the breeding process. In the case of storage proteins, our results didn't revealed any impact on the 12S globulin family.

#### *Comparison by shotgun proteomic method associated to iTRAQ labelling*

Using 2D electrophoresis, we were able to quantify and compare only one of the two major storage protein families in the seed, the 12S globulin. However some papers reported a variation of the ratio between 12S and 2S related to the breeding for glucosinolates. Therefore as napins (2S) were hardly detected due to their low molecular mass and the basicity of some isoforms, we set up an alternative shotgun proteomic method. The entire seed proteome of each variety is first hydrolysed, the resulting peptides are then labeled with one of the four isobaric tag and the derivatized digests are combined together. This mixture is then separated by a multi-dimensional chromatography prior mass spectrometry analysis.

**Table 1. Seed storage proteins relative quantification ( $\pm 0.2$ ) between the four varieties.**

	G/J (+/0+)	G/D (+/00)	J/D (0+/00)	D/Dn (00/00)
	117/116	117/115	116/114	115/114
12S family	0.84	0.81	0.92	1.09
2S family	1.04	1.53	1.66	1.01

This strategy resulted in 4000 spectra (three repetitions) from which 457 peptides were identified and quantified. Among these, 66 originated from storage proteins and permitted the identification of 13 different storage proteins. In the case of the 2S family, 7 different napins were identified, each by at least three specific peptides. Their relative quantification between the four varieties revealed that most of these proteins were less expressed in the low glucosinolates varieties. Another analysis of our results was also performed by grouping peptides of the same family and used all of them to quantify the family. The results obtained for 2S and 12S families are reported in Table 1. The relative quantification across the four lines revealed a relative stability of the expression of the 12S family given the maximum standard deviation of 20% (Ross et al., 2004). This was also the case when looking at the isoforms individually. On the contrary, the expression of the 2S decreased in the double null varieties lowered in their glucosinolates content.

## Conclusion

Numerous proteins differentially expressed have been identified and quantified. The two strategies developed here are very complementary for the type of proteins detected. Some proteins were identified by both methods; the measurement of their differential expression give very similar results and therefore reinforce our findings. The expression of napin decreased in varieties selected for their lower amount in glucosinolates while at the same time the expression of globulins remained stable. These results corroborate those on the increase of the 12S/2S ratio in double null varieties compared to their parents.

## References

- Renard, M.; Delourme, R.; Barret, P.; Brunel, D.; Froger, N.; Tanguy, X. (2001) Mutant gene of the gras family and plants with reduced development containing said mutant gene. WO 0109356, 2001.
- Raab, B.; Leman, H.; Schwenke, K. D.; Kozłowska, H. Comparative study of the protein patterns of some rapeseed (*Brassica napus* L.) varieties by means of polyacrylamide gel electrophoresis and high-performance liquid chromatography. *Nahrung* 1992, 36, (3), 239-247.
- Malabat, C.; Atterby, H.; Chaudhry, Q.; Renard, M.; Guéguen, J. Genetic variability of rapeseed protein composition. In 11 th International Rapeseed Congress. Towards Enhanced Value of Cruciferous Oilseed Crops by Optimal Production and use of the High Quality Seed Components, Copenhagen, DNK, 2003; Sorensen, H.; Sorensen, J. C.; Sorensen, S.; Bellostas Muguera, N.; Bjerregaard, C., Eds. pp 205-208.
- Albertin, W.; Brabant, P.; Catrice, O.; Eber, F.; Jenczewski, E.; Chevre, A. M.; Thiellement, H. Autopolyploidy in cabbage (*Brassica oleracea* L.) does not alter significantly the proteomes of green tissues. *Proteomics* 2005, 5, (8), 2131-9.
- Chevalier, F.; Martin, O.; Rofidal, V.; Devauchelle, A. D.; Barteau, S.; Sommerer, N.; Rossignol, M. Proteomic investigation of natural variation between *Arabidopsis* ecotypes. *Proteomics* 2004, 4, (5), 1372-81.
- Stotz, H. U.; Pittendrigh, B. R.; Kroymann, J.; Weniger, K.; Fritsche, J.; Bauke, A.; Mitchell-Olds, T. Induced plant defense responses against chewing insects. Ethylene signaling reduces resistance of *Arabidopsis* against Egyptian cotton worm but not diamondback moth. *Plant Physiol* 2000, 124, (3), 1007-18.
- Matsushima, R.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I. A novel ER-derived compartment, the ER body, selectively accumulates a beta-glucosidase with an ER-retention signal in *Arabidopsis*. *Plant J* 2003, 33, (3), 493-502.
- Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlett-Jones, M., He F., Jacobson, A., Pappin, D.J. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics*. 2004 Dec;3(12):1154-69