

Proteomic Variation Is as Large within as between Strawberry Varieties

Rikard Alm,^{*,†} Andreas Ekefjård,[‡] Morten Krogh,[§] Jari Häkkinen,[§] and Cecilia Emanuelsson[†]

Department of Biochemistry, Lund University, Lund, Sweden, Ludesi, Lund, Sweden, and Department of Theoretical Physics, Lund University, Lund, Sweden

Received January 25, 2007

In search for a strawberry (*Fragaria ananassa*) with low allergen content, we determined the proteomic variation within and between different varieties. Proteomics data were generated by DIGE and proteins identified with MALDI-MS/MS. The amount of the strawberry allergen Fra a 1 varied between different strawberry varieties (CV = 39%). The variation was at the same level, or even slightly larger, due to different growth conditions (CV = 43%). For 153 other proteins, the biological variation was more affected by different growth conditions than by different varieties (mean CV = 52% and 43%, respectively) due to variation in a subset of proteins. Thus, the allergen variation due to growth conditions must be taken into consideration in attempts to obtain a low-allergen strawberry. However, the allergen content was always lower in colorless (white) strawberry varieties than in the red ones. Moreover, of the spots whose expression correlated with the allergen and the color (32 and 68, respectively), only 3 were the same. This implies that these two phenotypic traits are not inseparable, and it may be possible to breed a red strawberry with low amount of allergen.

Keywords: proteomics • biological variation • variety • phenotypic plasticity • allergen • growth conditions • DIGE • external validity • intraspecific

Introduction

There is a biological variation between different varieties of cultivated plants. Variation between varieties is suitable for studying variation within a species (so-called intraspecific variation), and is the origin of plant species differentiation.¹ There are various techniques to estimate the variation among varieties, such as genotyping by DNA-based techniques like restriction fragment based polymorphism (RFLP) and random amplified polymorphic DNA (RAPD).² There are many studies that connect genetic variation to phenotype (reviewed in ref 1), usually morphological differences caused by loss of function or changed expression level in single genes. Genes related to adaptation to an environment are, on the other hand, rare or hard to detect,³ probably due to large phenotypic plasticity and a multigenic background. To estimate differences in phenotype and gene expression, strawberry microarray analyses have been performed, to yield information about up- and down-regulation of transcripts encoding enzymes tentatively important for firmness of berries^{4,5} or for biosynthesis of strawberry flavor compounds.^{6–9}

Proteomics is another powerful approach to investigate phenotype and gene expression differences between varieties. The change of protein abundances reflects the cellular response

to the environment, and some changes like post-translational modifications can only be seen on the protein level. Proteome analyses are limited by protein identification, which relies on the availability of sequence data. For plant species such as barley,¹⁰ oak,¹¹ banana,¹² or strawberry,¹³ with unsequenced genomes, proteomics can still be applied but gives a lower protein identification rate based largely on homologous sequences. Genome sequencing has been completed so far only for *Arabidopsis thaliana* and rice.

Fruit allergy is a growing problem in the western world. Many food companies use fruit as an additive to enhance taste and nutrition value.¹⁴ Since the major treatment for fruit allergy is avoidance and rescue medicine, low amounts of allergen in food is desirable. Fruit allergy is to a large extent depending on specific proteins that cross-react with human antibodies formed against homologous proteins in pollen from grass, birch, and other pollen producing plants.^{15–18} The strawberry allergen Fra a 1 is such a homologue to the major birch pollen allergen Bet v 1. Around 30% of patients in northern Europe with self reported food hypersensitivity also have adverse reactions to strawberry.¹⁹

We have previous shown that a “white”, colorless strawberry cultivar contained low amounts of the strawberry allergen Fra a 1.¹³ Such a low-allergen strawberry variety is however not available commercially as a registered variety, since the berries are of too low quality (small, soft, and not red). We have here extended the proteomic investigation to different commercially available strawberry varieties, to gain better insight into the variation in allergen content between the different varieties,

* To whom correspondence should be addressed. Department of Biochemistry, Center for Chemistry and Chemical Engineering, Lund University, P.O Box 124, S-221 00 Lund, Sweden. E-mail: Rikard.Alm@biochemistry.lu.se.

† Department of Biochemistry, Lund University.

‡ Ludesi.

§ Department of Theoretical Physics, Lund University.

as well as their proteome variation. Such information is important in understanding how a low allergenic red strawberry could be produced.

If technical replicates are used in an experiment, the variation in measurements can be divided into biological and technical variation. The biological variation can further be divided into genetic and environmental variation. Certified growers of strawberry propagate plants asexually from a nuclear meristem stock by runners (stolons) to enhance a certified variety for resale to producers. Since most professional strawberry producers buy strawberry plants from certified growers, this means that all plants from one variety, in principle, are genetically identical. Proteomic variation within a variety, therefore, represents adaptation of individual plants and berries to local differences in environment, so-called phenotypic plasticity.²⁰ Therefore, the biological variation between two varieties stems from both genetic differences between varieties, and from phenotypic plasticity due to environment.

To estimate the proteomic variation between varieties (intraspecific variation), various red strawberry varieties, representing a broad spectrum of different varieties that were not closely related, were randomly selected and grown under identical growth conditions. To estimate the proteomic variation within a variety, one red variety was collected from different growth conditions representing a realistic variation in growth conditions for commercial strawberry production in southern Sweden. We also included in this study more white strawberry cultivars, to investigate whether white strawberries are consistently low in allergen content. We found that approximately 10% of the proteins were differently expressed between red and white strawberries ($p < 0.01$), and that all three white strawberry cultivars were low in allergen. Between the red strawberry varieties, there was variation in allergen content, but none had an allergen content as low as the white ones. Moreover, the amount of strawberry allergen varied not only between the different varieties, but also between the different growth conditions. Also for 153 other strawberry proteins measured in our proteomics experiment, the biological variation was more affected by growth conditions than by differences between varieties.

Materials and Methods

Strawberry Material for Proteome Analysis. Two sets of strawberry samples were collected, as summarized in Figure 1. One set was collected to assess the biological variation between varieties, which were cultivated outdoors in June–July 2002, harvested at the same site, within a breeding program directed by Karin Trajkovski at Balsgård experimental station, Swedish University of Agricultural Sciences, Sweden. This set was composed of biological duplicates of four commercially available varieties of strawberry (Sengana, Zephyr, Cavendish, and Bogota) and two white strawberry cultivars (976902 and 945501). Additionally, to increase external validity, three more varieties were collected from another cultivation site (Höjebromölla, Lund, Sweden, owner Mårten Persson), two commercially available red varieties (Korona and Polka), and one white nameless and uncharacterized cultivar (04W). The different red varieties were generally not closely related, that is, neither parents nor grandparents were identical, with the exception that Sengana is one of the four grandparents to Korona. The ancestry of 04W is unknown, but it might have the same ancestry as the above-mentioned white cultivars (originating from Åke Truedsson, Klagshamn, Sweden). For

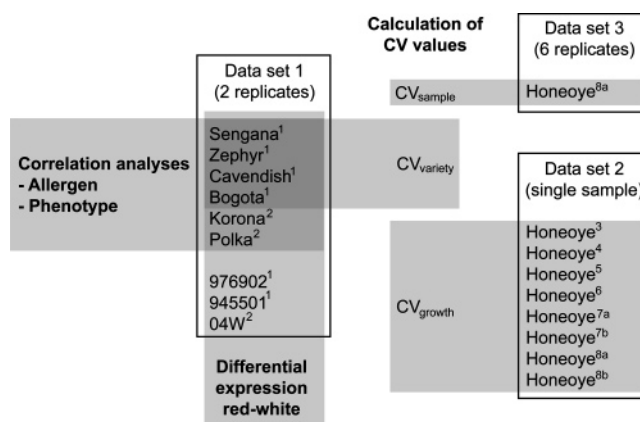


Figure 1. Overview of samples from different varieties and growth conditions used for statistical analyses. Data set 1 was used for differential expression analysis between red and white varieties, for correlation analyses, and for calculation of CV_{variety}. Data sets 2 and 3 were used for calculation of CV_{growth} and CV_{sample}, respectively. All samples are red strawberries except 976902, 945501, and 04W which were white. The superscript indicates different growth conditions: 1, Balsgård; 2, Höjebromölla; 3, Stenestad; 4, Genarp; 5, Hässleholm; 6, Sjöbo; 7a, Torsebro; 7b, Torsebro with cover; 8a, Rånna; and 8b, Rånna with cover.

simplicity, also the white strawberry cultivars are from here on designated “varieties”, although they are not registered as varieties.

The second set of strawberry samples was collected to assess the biological variation within a variety, with berries collected and harvested from different growth conditions representing a realistic variation in growth conditions for commercial strawberry production. This set was composed of berries from the commercially available strawberry variety, Honeoye, cultivated outdoors in June–July 2005 by different growers at six different locations in southern Sweden (Stenestad, Genarp, Hässleholm, Sjöbo, Torsebo, and Rånna experimental station, the Swedish University of Agricultural Sciences); from the two latter locations, berries were grown without and with a fabric or plastic cover. In total, single samples of berries, of the strawberry variety Honeoye, were collected from eight different growth conditions. In addition, to estimate sample variation (berry-to-berry variation), six replicate samples of the variety Honeoye were collected from one of the locations (Rånna).

1. Two-Dimensional Gel Electrophoresis. Proteins were extracted from individual strawberries (collected and harvested at ripe stage and stored until use at -20°C) by a phenol-extraction procedure as described previously.²¹ Separation of proteins by 2DE was performed at the Lund University Swegene Proteome Center (<http://www.swegene.se>) by DIGE,^{22,23} as described previously,¹³ but with IPG-strips (4–7 NL/24 cm) and with a Cy2-labeled pooled sample being the reference standard. Samples were labeled with the Cy-dyes (Cy3, Cy5) so that the replicates were dye-swapped. In white strawberries, the spot 23653 (the strawberry allergen Fra 1), was labeled with a false amplification in signal by Cy5. Spot volumes for spot 23653 from white strawberries derived from Cy5 were therefore excluded. The number of replicates from white varieties was therefore reduced from six to three, which is still sufficient for data analysis.

1.1. Image Analysis. The scanned 2DGE images were sent to Ludesi Analysis Center (Sweden, <http://www.ludesi.com>) for

image analysis using Ludesi's proprietary image analysis software. The protein spots were automatically detected, and the results were manually verified and edited where needed, with, on average, 1260 spots resolved by image analysis. The gels were automatically matched using all-to-all matching, avoiding introduction of bias caused by use of a reference gel. The matching was iteratively improved by optimization of matching parameters and manual editing. Integrated intensities were measured for each spot, background-corrected, and then normalized by mathematically minimizing the median expression difference between matched spots. The data set was also matched against the data set obtained in our previous study,¹³ in order to take advantage of proteins previously identified. Spots that after statistical analyses were found to correlate with the strawberry allergen, color, or other phenotypic traits were selected for spot picking and mass spectrometric protein identification as described below.

1.2. Preparative Gels. Each gel was loaded with 0.5 mg of protein and stained with CBB. Three preparative gels were run, each with proteins extracted from one of three strawberry varieties. The three varieties were chosen as to maximize the spot volumes of the proteins of interest. For each spot selected for spot picking, the three varieties with the highest value for the spot volume were given a score. All the scores were summarized, and the variety with the highest score was chosen as suitable for being run as preparative gel. This procedure was repeated using the spots with the lowest values for spot volumes in the chosen variety, such that for each protein selected for spot picking, it showed a good spot volume in at least one of the varieties finally selected. In total, 222, 182, and 158 spots were taken from three preparative gels with protein extracts from two red (Sengana and Polka) and one white cultivar (976902).

2. Statistical Analyses. Normalized spot volumes were used. Subsets of samples were selected for statistical analyses as outlined in Figure 1. Filtering of data were made such that a volume filter was applied, removing spots with volume <250 and spots with <80% presence in each subset.

2.1. Expression Analysis of Red and White Strawberries. The statistical test ANOVA was used to analyze the differential expression of proteins in red and white strawberries. With the above-mentioned criteria for filtering, in total, 526 spots were included in the expression analysis.

2.2. Proteins Correlating with Allergen. To find proteins whose expression correlated (positively or negatively) with the expression of the allergen, Kendall's rank correlation was used.²⁴ When the above-mentioned criteria for filtering was used, in total, 589 spots were included in the allergen correlation.

2.3. Proteins Correlating with Phenotype. Besides color, four other phenotypic traits were investigated; early ripening, flavor, firmness, and gray mould resistance. The different red strawberry varieties were compared, and each variety was assigned scores describing each phenotypic trait on a 5-graded scale. Scoring was made independently by two strawberry breeders, and mean values were calculated (see Table 1). The Jonckheere-Terpstra trend test²⁵ where used to find proteins whose expression correlated (positively or negatively) with a phenotypic score. When the above-mentioned criteria for filtering was used, in total, 589 spots were included in phenotype correlation.

2.4. Coefficient of Variation. A linear mixed model²⁶ was used for calculation of biological variation, expressed as coef-

Table 1. Scoring of Phenotypic Properties in Different Strawberry Varieties^a

	early ripening ^b	flavor ^c	firmness ^d	gray mould ^e
Sengana	4	2.75	2	5
Zephyr	1	3.5	2	2.5
Cavendish	2	4	3.5	3
Bogota	5	1	3	2
Polka	3.25	3.5	3.5	3
Korona	3	4.5	3	3

^a The red strawberry varieties were scored on a 5-graded scale with regard to phenotypic traits. Scoring was made independently by two strawberry breeders, and mean values were calculated. ^b Early ripening: 1 = early, 5 = late. ^c Flavor: 1 = low, 5 = high. ^d Firmness: 1 = soft, 5 = firm. ^e Gray mould: 1 = resistant, 5 = susceptible.

ficient of variation (CV). The model corrects for protein-specific dye effects. The model estimates both biological and technical variation, so the biological CVs calculated do not include technical variation. Here, technical variation means sample variation (berry-to-berry variation) combined with all the experimental variations. The linear mixed model is available as a Java program, named DIGEanalyzer, from (<http://bioinfo.thep.lu.se/digeanalyzer.html>). When the above-mentioned criteria for filtering was used, in total, 589 spots were included in the allergen correlation.

2.5. Estimation of Differential Variation. For each of 153 proteins, a Z-score was calculated as $(CV_{\text{variety}} - CV_{\text{growth}})/SD$, where SD is the standard deviation of the numerator, which is calculated as the square root of the sum of variances of the two CVs. The variance of the CV comes from the Fisher Information matrix of the linear mixed model and is calculated by DIGEanalyzer.²⁶ The Z-score is approximately normally distributed, and the *p*-value for a protein to have larger or smaller CV_{variety} than CV_{growth} is assuming normality. A Z-score with an absolute value larger than 3.1 is significant at the 0.001 level. At this level, 0.15 proteins are expected by chance.

3. Spot Picking and Mass Spectrometry for Protein Identification. Proteins correlating significantly ($p < 0.05$) with color, allergen, or phenotype were selected for spot picking and protein identification. Automated spot picking was performed as previously described.¹³ MALDI-MS and MS/MS mass spectra were acquired on a Applied Biosystems 4700 Proteomics Analyzer with TOF/TOF optics (Applied Biosystems, Darmstadt, Germany) in positive reflector mode. The following settings were used for MS acquisition: 3500 shots, 2800 laser intensity, and for MS/MS acquisition, 5500 shots, 3500 laser intensity, CID off. Peak extraction, calibration, and removal of contaminants were done by the 4700 Proteomic Analyzer built-in functions, and by an automatic procedure in PIUMS (<http://idelnx81.hh.se/bioinf/27,28>), a software designed for peak extraction and MS data analysis.

4. Protein Identification and Clustering of Mass Spectra.

4.1. Reversed Database. All database searches were performed against the taxonomic division "Plants" of the protein sequence database UniProt knowledgebase SwissProt/TrEMBL 20060312 with reversed sequences appended.²⁹ The script decoy.pl from Matrix Science (http://www.matrixscience.com/help/decoy_help.html) was used to create the reversed form of the database. The following parameters were used for a standard database search: Precursor tolerance, 20 ppm; MS/MS fragment tolerance, 0.2 Da; missed cleavages, 1; allowed modifications, carbamidomethyl on cysteine, acrylamide adduct (propionamide) on cysteine, and Oxidation of methionine, all as partial modifications.

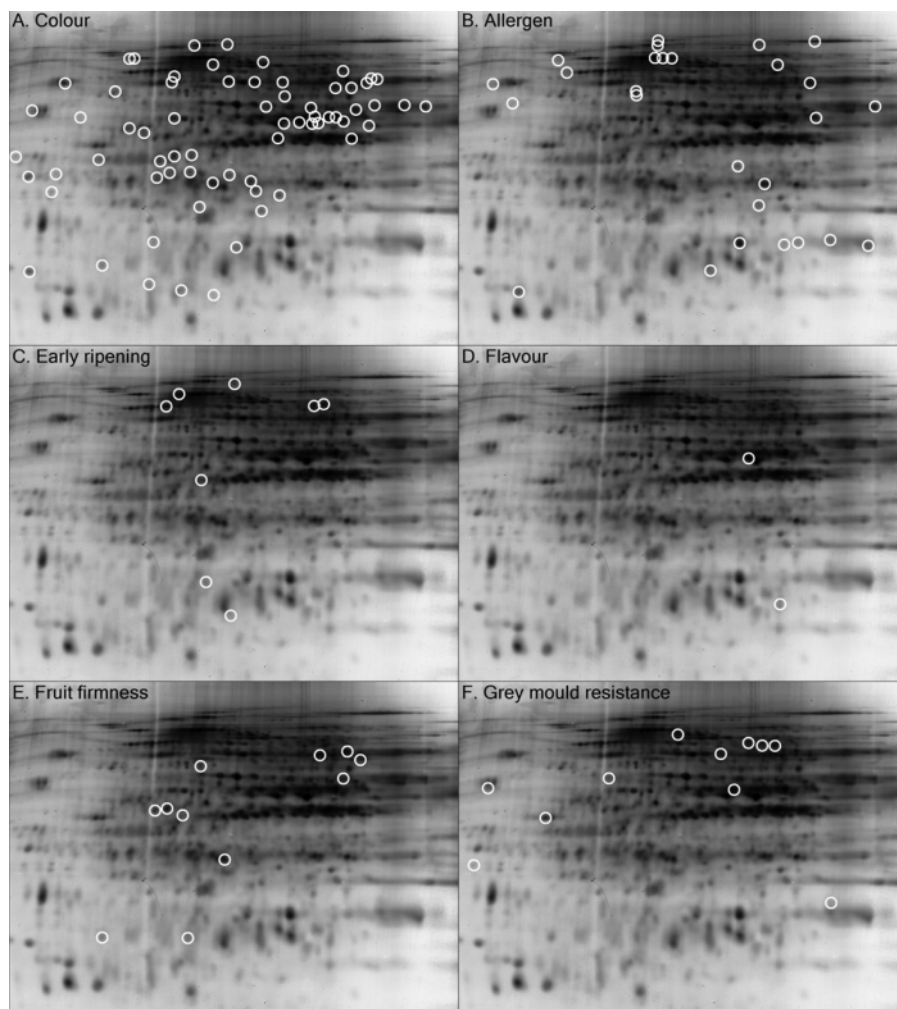


Figure 2. Proteins correlated with (A) color (68 of 526 spots), (B) allergen (32 of 589 spots), (C) early ripening (8 of 589 spots), (D) flavor (2 of 589 spots), (E) fruit firmness (11 of 589 spots), and (F) gray mould resistance (11 of 589 spots). Images of DIGE-2DE show proteins correlating with different phenotypic traits in statistical analyses of samples as outlined in Figure 1 and scoring according to Table 1. For protein identifications, see Table 2.

4.2. Protein Identification. Protein identification was performed with two search engines, Mascot³⁰ and PIUMS^{27,28} in the following way; (1) Protein mass fingerprint (pmf) search using MS-data and the software PIUMS; (2) MS/MS acquisition of the 5 most intense MS-peaks, followed by a combined MS–MS/MS search in Mascot; (3) MS/MS acquisition of MS-peaks selected by clustering (explained below), followed by a combined MS–MS/MS search in Mascot.

4.3. Criteria for Positive Identification. Protein identification in this study was done in two steps. First, suggested identifications with either a PIUMS protein score above 2.5 or Mascot protein score above 40 were retrieved. Second, from these suggested identifications, a match was deemed positive if either the protein was identified in at least 2 gels, or the protein was found by both search engines.³¹ The false-positive rate was calculated by counting the number of matches against reversed entries that matched the above-mentioned criteria and were found to be <1%. For proteins present in only one gel, single Mascot matches were accepted only if the protein score were above 1% false-positive level (corresponding to a protein score of 55).

4.4. Clustering of MS Spectra before MS/MS. With the software SPECLUST,³² (<http://bioinfo.thep.lu.se/speclust.html>),

peak lists from MS spectra from different spots were clustered, and peaks (peptides) shared within a cluster of mass spectra were used for selection of precursor masses to MS/MS analysis.

Results and Discussion

Proteome Analysis and Protein Identification. Samples from different strawberry varieties grown under different growth conditions were analyzed by 2D-DIGE, and subsets of samples were selected for statistical analyses as outlined in Figure 1. Of approximately 550 protein spots, 68 protein spots correlated with color (Figure 2A), and 32 proteins spots correlated with allergen (Figure 2B), and a few proteins correlated with other phenotypic traits (Figures 2C–F) as further described below.

For protein identification, 241 spots from different correlations were selected for mass spectrometry by spot picking from preparative gels (which on average resolved 849 spots), resulting in 97 proteins successfully identified as listed in Table 2. A success rate of protein identification of approximately 40% was obtained. Our approach, with identification criteria based on multiple preparative gels and two search engines³¹ used in combination with an appended²⁹ reversed database,^{33,34} allowed a lower than normal threshold for search-engine scores for

Table 2. Protein Identification and Variation in Strawberry

spot no.	Swiss-Prot accession no.	protein name	species	identification						variation			
				gels ^a	Piums ^b	Mascot ^c	peptide count ^d	total ion score ^e	best ion score ^f	CV _{variety} % ^g	CV _{growth} % ^h	Z-score ⁱ	corr ^j $p < 0.01$
23677	P49608	Aconitate hydratase, cytoplasmic	<i>C. maxima</i>	2	7.3	44	15	22	22				
22703	Q9XHH1	Actin	<i>M. pusilla</i>	2		155	15	84	48	11	31	−0.98	
22722	Q9XHH1	Actin	<i>M. pusilla</i>	3	17.0	165	17	84	48				
22830	Q7XZJ2	Actin	<i>G. hirsutum</i>	2	14.8	84	15	17	17	19	43	−0.42	
23066	Q9SMG4	Acyl-CoA binding protein (ACBP)	<i>D. lanata</i>	1		80	3	65	50	82	73	0.12	
22857	Q9AT16	Ankyrin-repeat protein HBP1	<i>N. tabacum</i>	1		83	4	77	77	15	17	−0.06	EtF4
22912	Q852H6	Anthranilate synthase beta subunit, Putative	<i>O. sativa</i>	1		154	2	154	86				A1
23552	Q7HJM1	ATP synthase alpha chain	<i>B. vulgaris</i>	2	7.6	41	18	15	15				
23410	Q01859	ATP synthase beta chain, mitochondrial precursor	<i>O. sativa</i>	1	4.8	91	19	61	39				
23710	P19023	ATP synthase beta chain, mitochondrial precursor	<i>Zea mays</i>	2	14.3	193	20	155	53	10	23	−2.85	
22729	O24345	ATP synthase, beta subunit, F1-ATP [Fragment]	<i>S. bicolor</i>	3	7.4	180	21	122	52	10	27	−2.54	A1
23292	O24345	ATP synthase, beta subunit, F1-ATP [Fragment]	<i>S. bicolor</i>	3	12.7	120	20	56	56				A1
23704	O24345	ATP synthase, beta subunit, F1-ATP [Fragment]	<i>S. bicolor</i>	1	3.3	74	18	31	18				
23237	O24346	ATP synthase, beta subunit, F1-ATP [Fragment]	<i>S. bicolor</i>	1	19.5	172	19	137	40	108	246	−0.23	
23608	Q84L45	CDC48-like protein [Fragment]	<i>M. jalapa</i>	1	6.0	45	7	22	22				
22785	Q4AE11	Chalcone isomerase	<i>F. ananassa</i>	2	5.0	215	10	173	62				A1
22866	Q4AE08	Chalcone synthase	<i>F. ananassa</i>	2	6.2	77	11	43	25.27				A1
23611	P30079	Chalcone synthase	<i>P. sylvestris</i>	1		75	11	59	59				A1
22781	O48919	Cytosolic ascorbate peroxidase	<i>F. ananassa</i>	3	8.8	144	10	94	57				B1
22786	Q9S7F7	Cytosolic ascorbate peroxidase APX19	<i>F. ananassa</i>	2	2.6	85	12	43	18				
22743	Q9S7F5	Cytosolic ascorbate peroxidase APX26	<i>F. ananassa</i>	3	7.7	146	14	88	46	17	29	−0.84	
23417	Q39575	Dynein gamma chain, flagellar outer arm	<i>C. reinhardtii</i>	2				46					
23303	P26301	Enolase 1	<i>Zea mays</i>	2	6.6	58	12	25	25	20	26	−0.47	
22704	Q66ME9	Flavanone 3-hydroxylase	<i>F. ananassa</i>	2	15.9	162	20	135	64	76	23	1.19	A1
22769	Q66ME9	Flavanone 3-hydroxylase	<i>F. ananassa</i>	2	5.7		21			81	13	0.62	A1
23690	Q66ME9	Flavanone 3-hydroxylase	<i>F. ananassa</i>	2	5.2	72	14	38	38	36	19	0.30	A1
23439	Q3T923	Fra a 1 allergen	<i>F. ananassa</i>	2	11.6	205	11	124	69				
23653	Q3T923	Fra a 1 allergen	<i>F. ananassa</i>	3	6.2	190	7	140	54				B
23864	Q3T923	Fra a 1 allergen	<i>F. ananassa</i>	2	7.8	113	8	66	40				
24556	Q3T923	Fra a 1 allergen	<i>F. ananassa</i>	2	14.1	180	12	97	45				
27277	Q3T923	Fra a 1 allergen	<i>F. ananassa</i>	2		73	11						
22753	Q7F445	Fructokinase, putative	<i>O. sativa</i>	2		95	6	87	37	13	38	−1.11	Et1
22769	O65918	GASA5-like protein [Fragment]	<i>P. mariana</i>	2		71	10			81	13	0.62	AtF4
24395	Q93VR3	GDP-mannose 3,5-epimerase	<i>A. thaliana</i>	2		50	8	32	32				
22711	Q944T7	Glyceraldehyde 3-phosphate dehydrogenase	<i>F. ananassa</i>	1		57	1	57	57				
22810	Q6YNS1	Glycine-rich RNA-binding protein	<i>P. avium</i>	3		96	5	84	84				
23096	Q6YNS1	Glycine-rich RNA-binding protein	<i>P. avium</i>	2		88	4	81	81	117	70	0.14	
23371	Q6YNS1	Glycine-rich RNA-binding protein	<i>P. avium</i>	2		81	3	73	73				
22853	Q02028	Heat shock 70 kDa chloroplast precursor	<i>P. sativum</i>	2	9.1	146	16	111	52				
22706	Q9SAU8	Heat shock 70 protein	<i>T. aestivum</i>	1		64	7	57	38				
22755	Q9SAU8	Heat shock 70 protein	<i>T. aestivum</i>	2	14.4	43	17						
22827	O50036	Heat shock 70 protein	<i>S. oleracea</i>	1	3.3	71	18	51	51	8	36	−1.28	A1
22975	Q9SAU8	Heat shock 70 protein	<i>T. aestivum</i>	1	11.1	42	16						
23144	Q9SAU8	Heat shock 70 protein	<i>T. aestivum</i>	1	11.6	44	17						
22817	Q8H6U4	Heat shock protein 60 [Fragment]	<i>P. dulcis</i>	2	10.6	54	15	46	38	22	38	−0.58	
23051	Q8H6U4	Heat shock protein 60 [Fragment]	<i>P. dulcis</i>	2	19.1	65	20			21	37	−0.97	
23133	O49962	Heat shock protein LMW	<i>F. ananassa</i>	3	7.3	170	8	162	57				
23094	Q2HSV5	Hypothetical protein	<i>M. truncatula</i>	2		51	6			24	109	−0.25	
22868	Q8H7F5	Hypothetical protein (Calcium ion binding)	<i>A. thaliana</i>	1		76	5	76	76				
22780	Q9M1G0	Hypothetical protein F9K21 (zinc-finger, metal-binding)	<i>A. thaliana</i>	2		53	16						
23027	O65501	Isocitrate dehydrogenase subunit 1, NAD+ dependent	<i>A. thaliana</i>	1		115	11	89	53				
23474	P52578	Isoflavone reductase homologue	<i>S. tuberosum</i>	1		77	4	69	69				
22791	Q08062	Malate dehydrogenase, cytoplasmic	<i>Zea mays</i>	2	12.4	118	11	85	62	15	25	−0.62	
23170	Q58FK4	Methylthiopentene dioxygenase 2 (1,2-dihydroxy-3-keto-5-)	<i>O. sativa</i>	2		69	4	62	62	92	146	−0.17	
22777	Q75K80	Nascent polypeptide associated complex	<i>O. sativa</i>	2	5.8	113	5	83	61				A1
22704	Q9M602	O-methyltransferase	<i>F. ananassa</i>	2	8.4	73	15			76	23	1.19	A1
23690	Q9M602	O-methyltransferase	<i>F. ananassa</i>	2	12.8	93	19	93	53	36	19	0.30	A1
22832	P26320	Oxygen-evolving enhancer protein 1, chloroplast pre	<i>S. tuberosum</i>	2	14.7	144	11	112	78				
22900	O22111	phosphogluconate dehydrogenase 6-	<i>G. max</i>	1	4.3	98	11	93	57				
23303	Q7XAS6	phosphoglycerate dehydrogenase 2-, Pollen	<i>C. dactylon</i>	2		55	8	55	30	20	26	−0.47	
22712	O24246	Phosphoglycerate mutase (2,3-bpg-independent)	<i>P. dulcis</i>	3	9.5	72	16	72	58	23	22	0.13	
22815	O24246	Phosphoglycerate mutase (2,3-bpg-independent)	<i>P. dulcis</i>	2	8.8	41	11	29	29	18	31	−0.45	
23181	Q40712	Polypeptide	<i>O. sativa</i>	2			27						A1
22735	P0C0Y3	Profilin Allergen Fra a 4	<i>F. ananassa</i>	1		102	4	88	61				
22834	P54776	Protease 26S, regulatory subunit 6A	<i>L. esculentum</i>	2	16.6	104	19	56	36	28	118	−0.79	
22807	P54778	Protease 26S, regulatory subunit 6B	<i>S. tuberosum</i>	3	17.7	130	20	81	39	30	62	−0.40	
22897	P54778	Protease 26S, regulatory subunit 6B	<i>S. tuberosum</i>	2	6.5		17			19	26	−0.59	
23646	Q8W3N9	Proteasome 26S, regulatory triple-A ATPase subunit3	<i>O. sativa</i>	2	27.2	72	21	21	21	21	27	−0.34	
22917	Q93 ' 33	proteasome subunit [Fragment] beta 3	<i>N. tabacum</i>	1	3.3	40	6	20	20				A1
22833	Q9M4T8	Proteasome subunit alpha type 5	<i>G. max</i>	2	6.7	93	11	50	30	22	14	1.03	
22702	Q94110	Putative quinone oxidoreductase	<i>F. ananassa</i>	3	12.0	90	17						
23862	Q94110	Putative quinone oxidoreductase	<i>F. ananassa</i>	1		172	3	165	70	97	34	0.44	A1
22728	Q94110	Putative quinone oxidoreductase	<i>F. ananassa</i>	3	10.4	200	18	134	59	32	31	0.04	A1D1
23764	Q94110	Putative quinone oxidoreductase	<i>F. ananassa</i>	2	16.7	146	14	71	57				A1
22751	O23979	Pyrophosphatase, soluble inorganic	<i>H. vulgare</i>	2		65	4	57	57	9	11	−0.87	Et1
23697	Q84V25	Quinone oxidoreductase	<i>F. ananassa</i>	2	6.9	116	14	82	44				
23702	Q84V25	Quinone oxidoreductase	<i>F. ananassa</i>	2	5.5	87	16	14	14				F1
22803	Q9STA6	RAD23 protein	<i>L. esculentum</i>	1		86	3	77	50	15	18	−0.67	
22941	O23942	Ripening-induced protein	<i>F. vesca</i>	1	11.5	133	10	86	42	140	83	0.28	AtEt1

Table 2. (Continued)

spot no.	Swiss-Prot accession no.	protein name	species	identification						variation			
				gels ^a	Piums ^b	Mascot ^c	peptide count ^d	total ion score ^e	best ion score ^f	CV _{variety} % ^g	CV _{growth} % ^h	Z-score ⁱ	corr ^j $p < 0.01$
23405	Q9ATC0	Rubisco activase	<i>Z. aethiopica</i>	1	13.6	42	13						F†
22745	P21239	RuBisCO large subunit-binding protein	<i>B. napus</i>	2	9.3	85	12	74	48	12	20	−0.84	F†
22841	Q9AT55	S-adenosyl-L-methionine synthetase	<i>E. umbellata</i>	3	9.1	154	14	109	48				E†
23441	Q8GTL5	S-adenosylmethionine synthase	<i>Carica papaya</i>	2	6.8	139	13	112	45				A†
22829	Q7M1T2	Seed storage protein Lec2SA1 [Fragment]	<i>L. esculentum</i>	2		62	6						
22780	Q8LAV0	Succinyl-CoA ligase beta subunit	<i>A. thaliana</i>	2		120	8	106	55				
22834	Q9AYS9	Tat binding protein like protein	<i>B. campestris</i>	2	20.1	99	20	43	26	28	118	−0.79	F†
23015	Q8GUR9	Thioredoxin h	<i>P. sativum</i>	3		93	3	88	88	42	23	0.67	A†
22856	Q9M4S8	Triosephosphate isomerase, chloroplast pre	<i>F. ananassa</i>	2	10.7	60	12	60	23				A†
22967	Q546I4	Tubulin Alpha	<i>Z. mays</i>	2	9.7		9						
23610	Q6Y1F1	Tubulin alpha chain	<i>T. repens</i>	1	17.9	86	14						
22778	Q93Y15	Tubulin Beta	<i>A. thaliana</i>	1	10.8	93	18	34	34	16	78	−1.28	A†
23059	Q9ZRB0	Tubulin beta-3 chain	<i>T. aestivum</i>	2	5.6	45	12			11	20	−0.43	
22916	Q93Y15	Tubulin, beta	<i>A. thaliana</i>	3	9.5	46	12			36	86	−0.28	
23035	Q6TYR6	Tubulin, beta 2	<i>P. patens</i>	2	9.1		12						
22747	Q4TYY6	Ubiquitinating enzyme	<i>A. thaliana</i>	3		43	5	36	36	16	24	−0.72	E†
23070	Q547B9	Ubiquitin-like modifier 1, small	<i>A. thaliana</i>	1	5.5	59	5	26	26	14	27	−0.75	A†
23206	Q651T8	Vacuolar proton-ATPase	<i>O. sativa</i>	2	15.8	47	19			27	37	−0.43	

^a The number of preparative gels the protein has been identified in (1–3). ^b Score for the search engine PIUMS (quality). ^c Score for the search engine Mascot (protein score, combining MS and MS/MS). ^d Number of peptides that match identified protein in MS. ^e Cumulative ion score for MS/MS in Mascot. ^f Best ion score for MS/MS in Mascot. ^g Coefficient of variation due to variety. ^h Coefficient of variation due to growth condition. ⁱ Test score showing if CV_{variety} is significantly different from CV_{growth}. A value of ± 1.96 is significant on $p < 0.01$ level. Positive values indicate that CV_{variety} is significantly larger than CV_{growth} and v.v. ^j Correlation and differential expression: A, color; B, allergen; C, early ripening; D, flavor; E, fruit firmness; F, gray mould resistance. Arrow indicate direction of correlation

maximum number of identifications without compromising the false-positive rate of 1%.

Proteins Down-Regulated in White Strawberries. In total, 68 protein spots were differently expressed ($p < 0.01$) between six red varieties and three white varieties (Figure 2A). Of these 68 spots, 24 different proteins were identified (Table 2) in 22 spots (some spots containing more than one protein). Compared to a previous study, where only two varieties (one red, one white) were compared,¹³ fewer protein spots (68, compared to previously 104) were found to be differentially expressed with respect to the color difference. This is due to the larger biological diversity among the nine different varieties used here, making proteins correlating with the color stand out more clearly. When only two varieties are compared, not only differences due to the color, but also differences between the two specific varieties are detected (lower external validity). Also confirmed here, as previously observed,¹³ is that enzymes in the pathway of flavonoid biosynthesis (e.g., chalcone isomerase, chalcone synthase, flavanone 3-hydroxylase, and O-methyl-transferase) were down-regulated in the white strawberries, and enzymes involved in protection against oxidative stress such as, for example, thioredoxin, were up-regulated. The oxidative stress is probably related to the loss of color and flavonoid synthesis. Since flavonoids protect against UV-radiation, a lower amount of flavonoids should lead to more damaged proteins and more oxidative stress.

Proteins Correlated with Allergen and Other Phenotypic Traits. The strawberry allergen is detected in five different spots, presumably differing in terms of post-translational modifications.³⁵ Of these five spots, one is however predominating (spot number 23653); the spot volume of this spot was therefore used for correlation analyses. By the use of Kendall's rank correlation on the data set composed of six different red strawberry varieties as outlined in Figure 1, protein spots were recognized whose expression correlated significantly ($p < 0.01$) with the expression of the allergen (Figure 2B), in total 32 spots. Only 3 (unidentified) of these 32 spots correlating with allergen were the same as the spots correlating with color (Figure 2A), implicating that these two phenotypic traits are not inseparable.

It may, therefore, be possible to separate the two traits by breeding, that is, to find a red strawberry with low amount of allergen.

To get a hint of the function of the strawberry allergen, whose biological function is not clear, it would be interesting to know the identity of the proteins correlating with the allergen. Unfortunately, protein identification was only possible for a few of these 32 proteins (Table 2) correlated at $p < 0.01$. A survey was made among the 105 proteins that correlated at $p < 0.05$ with the expression of the allergen. Among these, 10 proteins were identified. Although the expected number of false positives in this correlation is high (30%), their protein identifications (Q3T923 Fra a 1 allergen, Q9AU11 Aromatic polyketide synthase, Q4AE08 Chalcone synthase, O48919 Ascorbate peroxidase, Q02028 Heat shock protein70, Q8H6U4 Heat shock protein60, Q8H7F5 Hypothetical calcium binding, O22111 6-phosphogluconate dehydrogenase, Q7M1T2 Seed storage protein, Q4TYY6 Ubiquitinating enzyme) were all dissimilar from the identified proteins correlating with color, except for the Chalcone synthase (Q4AE08). Of the 10 above-mentioned identified proteins correlating with the allergen, two were heat shock proteins, perhaps reflecting the fact that the strawberry allergen homologues are stress-induced so-called PR-10 proteins.¹⁶

We also attempted to detect proteins correlating with the other phenotypic traits (varieties scored with respect to these traits according to Table 1), using the Jonckheere-Terpstra trend test. The number of proteins whose expression correlated significantly ($p < 0.01$) with early ripening, flavor, fruit firmness, and gray mould resistance were 11, 2, 8, and 11, respectively (Figure 2C–F), forming entirely different protein spot patterns (“proteome fingerprints”) than for color (2A) and allergen (2B). However, the false-positive rate is, on average, 50% (6 proteins); therefore, the spot pattern profiles in Figure 2C–F are not reliable as proteome fingerprints. Nevertheless, it is likely that similar proteome fingerprints could be improved and used as selection markers for specific traits by including a higher number of samples and varieties with more extreme scores (very high or very low) for a phenotypic trait. The varieties used

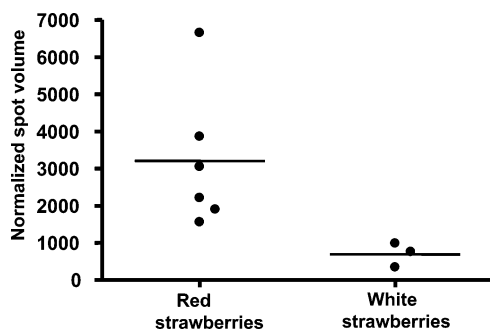


Figure 3. Lower abundance of the strawberry allergen Fra a 1 in white strawberries. Data obtained from differential expression analysis by 2D-DIGE (spot 23653, containing the strawberry allergen Fra a 1) showing spot volumes measured with Cy3 for red and white varieties (data set 1, Figure 1).

here only had very medium score values for these phenotypic traits (Table 1), for example, for fruit firmness ranging from 2 to 3.5. Proteome profiling, thus, needs to be further addressed using other strawberry varieties with more extreme scores (e.g., 1 or 5) with respect to these phenotypic traits. This was not the case here, since the protein profiling was in first place designed to investigate allergenicity in randomly selected red and white varieties.

A comprehensive interpretation of the current fingerprints was also hampered because protein identification was possible only for a minor part of the correlating proteins (the ones listed in Table 2). The low identification rate is partly due to the DIGE-technique and the advanced image analysis which allowed detection of many very weak spots which are more difficult in terms of protein identification. Although spectral quality was not bad, the protein identification does become dependent on a very good sequence coverage when identification has to rely on identification on homologous sequences due to lack of strawberry sequence data.

Allergen Content Is Low in All White Strawberry Varieties.

In the present study, with 6 different red and 3 white varieties, the external validity is increased compared to our previous comparison of allergen content in one colorless (white) variety and one red variety, Korona.¹³ Here, all the 3 white strawberry varieties investigated showed a significantly lower abundance of the allergen than any of the red varieties ($p = 0.021$) (Figure 3). It brings support to the idea that the allergen is involved in the synthesis of the strawberry color substance, maybe as a carrier of hydrophobic compounds,³⁶ and further suggests that any of the white varieties containing low amounts of allergen may be tolerated by people allergic to strawberry. Apple cultivars with low allergen content appear less allergenic.³⁸ A clinical study is therefore currently performed in which preliminary data suggest that white strawberries indeed do not elicit any allergic response (vanOdijk, J. and Bengtsson, U., Sahlgrenska University Hospital, Gothenburg, Sweden, personal communication).

Although none of the red varieties had as low amount of the allergen as the white ones, there was a 3.4-fold difference in the amount of allergen between the highest and lowest red strawberry variety (Figure 4). The duplicate values for the amount of allergen for some of the varieties were very similar (Zephyr, Polka), while for some of the varieties, the duplicate values varied more (Korona and Bogota). There is obviously variation not only between varieties, but also within varieties.

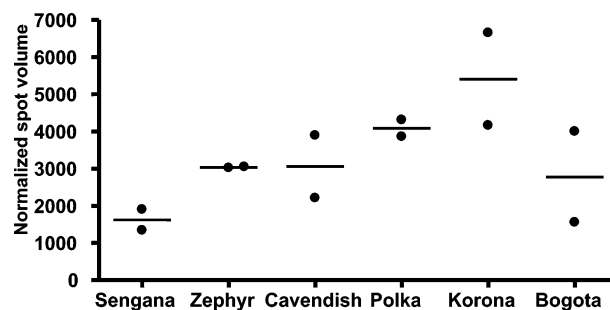


Figure 4. Variation in abundance of the strawberry allergen Fra a 1 depending on variety. Data obtained from 2D-DIGE (spot 23653, containing the strawberry allergen Fra a 1) showing spot volumes for six red varieties (from data set 1, Figure 1).

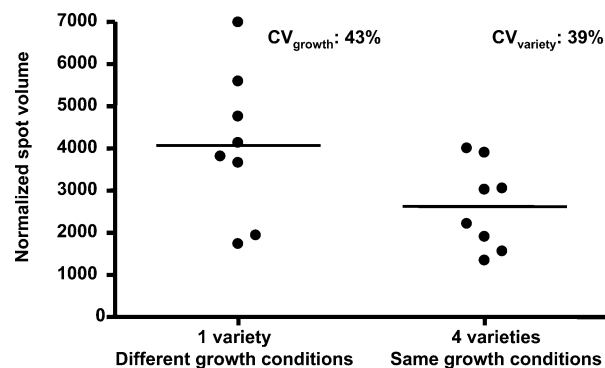


Figure 5. Variation in abundance of the strawberry allergen Fra a 1 due to different growth conditions as compared to different varieties. Data obtained from 2D-DIGE (spot 23653, containing the strawberry allergen Fra a 1) showing spot volumes for one variety grown under different growth conditions (data set 2, Figure 1) and different varieties grown under same growth conditions (upper four in data set 1, Figure 1)

Allergen Variation Is as Large within a Variety as between Varieties. As compared to the variation between varieties, we also analyzed the variation within one variety, Honeoye, collected from different growth conditions. Under these different growth conditions, variation in the amount of allergen within this variety was found to be more or less equal to the variation between varieties. The coefficients of variation (CV), that is, the relative standard deviation values, were similar or slightly larger within the variety (CV_{growth} 43%), compared to between the varieties (CV_{variety} 39%) as shown in Figure 5. The variation within berries grown at the same location, representing both technical and berry-to-berry variation, was lower (24%) than both CV_{growth} and CV_{variety} .

Thus, the amount of the strawberry allergen was affected as much by the growth conditions as by the genetic differences between the varieties. For apple, it has been documented that the apple allergen level may vary from year to year,³⁸ or due to storage conditions.³⁹ Taken together, this means that to obtain a strawberry with low amounts of allergen it will thus be important not only to compare varieties, but to evaluate the amount of allergen under many different conditions.

Proteome Variation Is Larger within a Variety Than between Varieties. To estimate the proteome variation also for other proteins than the strawberry allergen, the CV was calculated for the 153 proteins present in both data sets with at least 80% presence (Figure 6), using a linear mixed model.²⁶ When information from, in total, 5 red varieties and 9 different

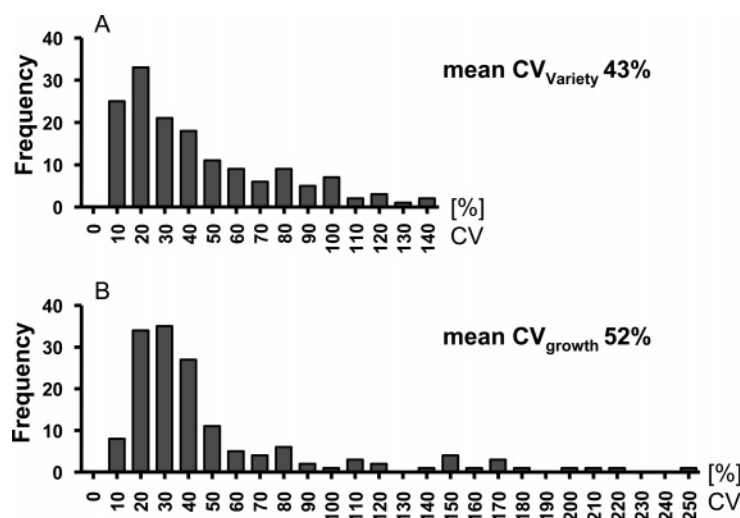


Figure 6. Proteome variation due to different varieties (A) and different growth conditions (B) using coefficient of variation (CV). The CV values were calculated by a linear mixed model for spot volumes of 153 proteins using eight replicates each, designated CV_{variety} and CV_{growth} , respectively, in Figure 1. A high number indicates a high variation.

growth conditions was used, the model calculated, for each of 153 proteins, a CV value for the variation due to variety (CV_{variety} , variation between varieties) and a CV value for the variation due to growth condition (CV_{growth} , variation within a variety), as listed in the columns to the right in Table 2. The average of these CV values was then calculated, and found to be 43% for CV_{variety} and 52% for CV_{growth} , as indicated in Figure 6. The average value for CV_{sample} was 37%.

Thus, the proteome variation was higher due to the growth conditions than due to the strawberry variety, but also quite high for berry-to-berry variation. The CV_{sample} value contained both the berry-to-berry variation and the technical variation which is usually around 25%⁴¹ and, in our case, estimated to 21% using the same linear mixed model calculation on other data sets.²⁶ This relatively large biological variation between individual berries was detectable because single strawberries, not pooled samples with several berries, were analyzed.

Data presented for oak^{11,41} are similar to our findings, with proteome variation for samples derived from different growth conditions with CV values around 59%, and proteome variation for different provenances grown under identical growth conditions with CV values around 40%. To estimate a minimal biological variation, a lower variation (CV 24%) was deliberately created in a study using *Medicago truncatula*⁴³ using a single seed line for high degree of genetic similarity, identical growth conditions, and pooling of 10 leaves per sample.

The higher variation due to differences in growth conditions compared to variety differences is not due to a higher variation in all proteins. As reflected by the distribution of the variation due to growth conditions, showing a longer tail to the right in Figure 6B compared to Figure 6A, there is a subset of proteins with higher CV values. In GFP-tagged *Saccharomyces cerevisiae* strains, a subset of proteins responding to changes in the environment, like those involved in stress-response, amino acid biosynthesis, and in heat shock were found to show high variation, whereas proteins involved in translational initiation, ribosomal proteins, and protein degradation exhibited low variation.⁴³ In our case, too few proteins have been identified for a comprehensive comparison of proteins with high and low CV_{growth} values.

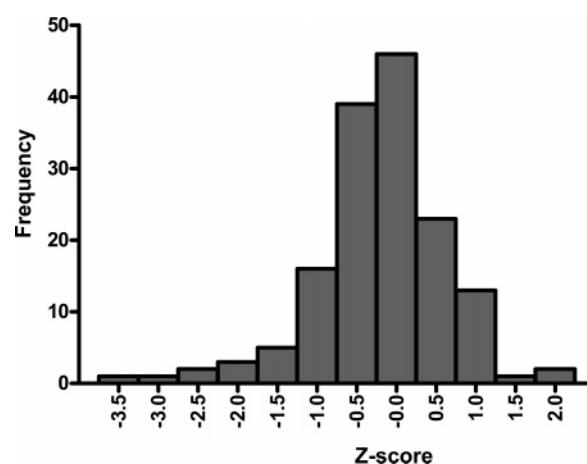


Figure 7. Distribution of pairwise differences between the CV_{variety} and CV_{growth} normalized by the standard deviation of this difference (Z-score). Values for CV_{variety} and CV_{growth} calculated as described in Figure 6. Cutoff scores for significance level $p = 0.001$ are +3.1 and -3.1. Proteins with negative values denote larger variation due to growth conditions. Positive values denote larger variation due to varieties.

Subsets of Proteins More Variable Due to Growth Conditions. To check if the difference in variation between CV_{variety} and CV_{growth} is reflected on the level of individual proteins, the two CV values of each protein were compared with a Z-test (Table 2, column designated Z-score, next most to the right). A positive Z-score denotes a higher variation due to variety than due to growth conditions, and a negative denotes a higher variation due to growth conditions than due to variety. The distribution of Z-scores is clearly skewed against lower Z-scores (Figure 7), giving further support to the notion that proteomic variation is larger due to growth conditions than due to variety.

Conclusions

Compared to the proteomal variation between strawberry varieties (so-called intraspecific variation¹), the variation due to the capacity of a strawberry variety to express different proteomal phenotypes depending on the growth conditions

(so-called phenotypic plasticity²⁰) was equally large. The different growth conditions here analyzed reflected a realistic variation in growth conditions to be considered for commercial strawberry cultivation. However, the allergen content was considerably lower in colorless (white) strawberry varieties under all the growth conditions tested. Moreover, very few of the spots whose expression correlated with the allergen and the color, respectively, were the same. This implies that these two phenotypic traits are not inseparable, and it may be possible to breed a red strawberry with consistently low amount of allergen.

Acknowledgment. Karin Trajkovski is acknowledged for contributing her experience of many years of strawberry plant breeding and for collecting strawberries. We thank the various strawberry growers for donating material and Birgitta Svensson at Rånna experimental station, the Swedish University of Agricultural Sciences, for strawberries and helpful advice. This study was supported by a grant from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning Dnr 225-2004-1790. M.K. is supported by the Swedish Foundation for Strategic Research, the Knut and Alice Wallenberg Foundation (SSF) Create Health Centre.

References

- Alonso-Blanco, C.; Mendez-Vigo, B.; Koornneef, M. From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. *Int. J. Dev. Biol.* **2005**, *49* (5–6), 717–732.
- Kaczmarek, E.; Hortyn'ski, J. The application of RAPD markers for genetic studies on *Fragaria x Fragaria* duch. (strawberry) with due consideration for permanent flowering. *Electron. J. Pol. Agric. Univ.* **2002**, *5*, (1).
- Kellogg, E. A. Evolution of developmental traits. *Curr. Opin. Plant Biol.* **2004**, *7* (1), 92–98.
- Jimenez-Bermudez, S.; Redondo-Nevado, J.; Munoz-Blanco, J.; Caballero, J. L.; Lopez-Aranda, J. M.; Valpuesta, V.; Pliego-Alfaro, F.; Quesada, M. A.; Mercado, J. A. Manipulation of Strawberry Fruit Softening by Antisense Expression of a Pectate Lyase Gene. *Plant Physiol.* **2002**, *128*, 751–759.
- Salentijn, E. M.; Aharoni, A.; Schaart, J. G.; Boone, M. J.; Krens, F. A. Differential gene expression analysis of strawberry cultivars that differ in fruit-firmness. *Physiol. Plant.* **2003**, *118*, 571–578.
- Aharoni, A.; Keizer, L. C.; Bouwmeester, H. J.; Sun, Z.; Alvarez-Huerta, M.; Verhoeven, H. A.; Blaas, J.; van Houwelingen, A. M.; De Vos, R. C.; van der Voet, H.; Jansen, R. C.; Guis, M.; Mol, J.; Davis, R. W.; Schena, M.; van Tunen, A. J.; O'Connell, A. P. Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. *Plant Cell* **2000**, *12* (5), 647–662.
- Aharoni, A.; Giri, A. P.; Verstappen, F. W.; Berteaux, C. M.; Sevenier, R.; Sun, Z.; Jongsma, M. A.; Schwab, W.; Bouwmeester, H. J. Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. *Plant Cell* **2004**, *16* (11), 3110–3131.
- Raab, T.; Lopez-Raez, J. A.; Klein, D.; Caballero, J. L.; Moyano, E.; Schwab, W.; Munoz-Blanco, J. FaQR, required for the biosynthesis of the strawberry flavor compound 4-hydroxy-2,5-dimethyl-3(2H)-furanone, encodes an enone oxidoreductase. *Plant Cell* **2006**, *18* (4), 1023–1037.
- Wein, M.; Lavid, N.; Lunkenbein, S.; Lewinsohn, E.; Schwab, W.; Kaldenhoff, R. Isolation, cloning and expression of a multifunctional O-methyltransferase capable of forming 2,5-dimethyl-4-methoxy-3(2H)-furanone, one of the key aroma compounds in strawberry fruits. *Plant J.* **2002**, *31* (6), 755–765.
- Ostergaard, O.; Melchior, S.; Roepstorff, P.; Svensson, B. Initial proteome analysis of mature barley seeds and malt. *Proteomics* **2002**, *2* (6), 733–739.
- Jorge, I.; Navarro, R. M.; Lenz, C.; Ariza, D.; Jorrián, J. Variation in the holm oak leaf proteome at different plant developmental stages, between provenances and in response to drought stress. *Proteomics* **2006**, *6* (Suppl. 1), S207–214.
- Carpentier, S.; Witters, E.; Laukens, K.; Deckers, P.; Swennen, R.; Panis, B. Preparation of protein extracts from recalcitrant plant tissues: An evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* **2005**, *5* (10), 2497–2507.
- Hjerno, K.; Alm, R.; Canback, B.; Matthiesen, R.; Trajkovski, K.; Bjork, L.; Roepstorff, P.; Emanuelsson, C. S. Down-regulation of the strawberry Bet v 1-homologous allergen in concert with the flavonoid biosynthesis pathway in colourless strawberry mutant. *Proteomics* **2006**, *6* (5), 1574–1587.
- Hannum, S. M. Potential impact of strawberries on human health: a review of the science. *Crit. Rev. Food Sci. Nutr.* **2004**, *44* (1), 1–17.
- Vieths, S.; Scheurer, S.; Ballmer-Weber, B. Current understanding of cross-reactivity of food allergens and pollen. *Ann. N. Y. Acad. Sci.* **2002**, *964*, 47–68.
- Aalberse, R. C.; Akkerdaas, J.; van Ree, R. Cross-reactivity of IgE antibodies to allergens. *Allergy* **2001**, *56* (6), 478–490.
- Breiteneder, H.; Ebner, C. Molecular and biochemical classification of plant-derived food allergens. *J. Allergy Clin. Immunol.* **2000**, *106* (1 Pt 1), 27–36.
- Mills, E.; Madsen, C.; Shewry, P.; Wichers, H. Food allergens of plant origin - their molecular and evolutionary relationships. *Trends Food Sci. Technol.* **2003**, *14*, 145–156.
- Eriksson, N. E.; Moller, C.; Werner, S.; Magnusson, J.; Bengtsson, U.; Zolubas, M. Self-reported food hypersensitivity in Sweden, Denmark, Estonia, Lithuania, and Russia. *J. Invest. Allergol. Clin. Immunol.* **2004**, *14* (1), 70–79.
- Sultan, S. E. Phenotypic plasticity for plant development, function and life history. *Trends Plant Sci.* **2000**, *5* (12), 537–542.
- Karlsson, A. L.; Alm, R.; Ekstrand, B.; Fjellkner-Modig, S.; Schiott, A.; Bengtsson, U.; Bjork, L.; Hjerno, K.; Roepstorff, P.; Emanuelsson, C. S. Bet v 1 homologues in strawberry identified as IgE-binding proteins and presumptive allergens. *Allergy* **2004**, *59* (12), 1277–1284.
- Lilley, K. S.; Friedman, D. B. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Rev. Proteomics* **2004**, *1* (4), 401–409.
- Wu, T. L. Two-dimensional difference gel electrophoresis. *Methods Mol. Biol.* **2006**, *328*, 71–95.
- Kendall, M. G. A New Measure of Rank Correlation. *Biometrika* **1938**, *30* (1–2), 81–93.
- Jonckheere, A. A distribution-free K sample test against ordered alternatives. *Biometrika* **1954**, *41*, 133–145.
- Krogh, M.; Liu, Y.; Bengtsson, S.; Valastro, B.; James, P. Analysis of DIGE data using a linear mixed model allowing for protein-specific dye effects. *J. Proteome Res.* **2006**, submitted for publication.
- Levander, F.; Rognvaldsson, T.; Samuelsson, J.; James, P. Automated methods for improved protein identification by peptide mass fingerprinting. *Proteomics* **2004**, *4* (9), 2594–2601.
- Samuelsson, J.; Dalevi, D.; Levander, F.; Rognvaldsson, T. Modular, scriptable and automated analysis tools for high-throughput peptide mass fingerprinting. *Bioinformatics* **2004**, *20* (18), 3628–3635.
- Higdon, R.; Hogan, J. M.; Belle, G. V.; Kolker, E. Randomized sequence databases for tandem mass spectrometry peptide and protein identification. *OMICS: J. Integr. Biol.* **2005**, *9* (4), 364–379.
- Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551–3567.
- Levander, F.; Krogh, M.; Warell, K.; Garden, P.; James, P.; Hakkinen, J. Automated reporting from gel-based proteomics experiments using the open source Proteios database application. *Proteomics* **2007**, *7* (5), 668–674.
- Alm, R.; Johansson, P.; Hjerno, K.; Emanuelsson, C.; Ringner, M.; Hakkinen, J. Detection and identification of protein isoforms using cluster analysis of MALDI-MS mass spectra. *J. Proteome Res.* **2006**, *5* (4), 785–792.
- Moore, R. E.; Young, M. K.; Lee, T. D. Qscore: an algorithm for evaluating SEQUEST database search results. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (4), 378–386.
- Resing, K. A.; Meyer-Arendt, K.; Mendoza, A. M.; Aveline-Wolf, L. D.; Jonscher, K. R.; Pierce, K. G.; Old, W. M.; Cheung, H. T.; Russell, S.; Wattawa, J. L.; Goehle, G. R.; Knight, R. D.; Ahn, N. G. Improving reproducibility and sensitivity in identifying human proteins by shotgun proteomics. *Anal. Chem.* **2004**, *76* (13), 3556–3568.

- (35) Musidlowska-Persson, A.; Alm, R.; Emanuelsson, C. Cloning and sequencing of the Bet v 1-homologous allergen Fra a 1 in strawberry (*Fragaria ananassa*) shows the presence of an intron and little variability in amino acid sequence. *Mol. Immunol.* **2007**, *44* (6), 1245–1252.
- (36) Mogensen, J. E.; Wimmer, R.; Larsen, J. N.; Spangfort, M. D.; Otzen, D. E. The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. *J. Biol. Chem.* **2002**, *277* (26), 23684–23692.
- (37) Bolhaar, S. T.; van de Weg, W. E.; van Ree, R.; Gonzalez-Mancebo, E.; Zuidmeer, L.; Bruijnzeel-Koomen, C. A.; Fernandez-Rivas, M.; Jansen, J.; Hoffmann-Sommergruber, K.; Knulst, A. C.; Gilissen, L. J. In vivo assessment with prick-to-prick testing and double-blind, placebo-controlled food challenge of allergenicity of apple cultivars. *J. Allergy Clin. Immunol.* **2005**, *116* (5), 1080–1086.
- (38) Marzban, G.; Puehringer, H.; Dey, R.; Brynda, S.; Ma, Y.; Martinelli, A.; Zaccarini, M.; van der Weg, E.; Housley, Z.; Kolarich, D. Localisation and distribution of the major allergens in apple fruits. *Plant Sci.* **2005**, *169* (2), 387–394.
- (39) Sancho, A. I.; Foxall, R.; Browne, T.; Dey, R.; Zuidmeer, L.; Marzban, G.; Waldron, K. W.; vanRee, R.; Hoffmann-Sommergruber, K.; Laimer, M.; Mills, E. N. C. Effect of postharvest storage on the expression of the apple allergen Mal d 1. *J. Agric. Food Chem.* **2006**, *54* (16), 5917–5923.
- (40) Molloy, M. P.; Brzezinski, E. E.; Hang, J.; McDowell, M. T.; VanBogelen, R. A. Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics* **2003**, *3* (10), 1912–1919.
- (41) Jorge, I.; Navarro, R. M.; Lenz, C.; Ariza, D.; Porras, C.; Jorin, J. The holm oak leaf proteome: analytical and biological variability in the protein expression level assessed by 2-DE and protein identification tandem mass spectrometry de novo sequencing and sequence similarity searching. *Proteomics* **2005**, *5* (1), 222–234.
- (42) Asirvatham, V. S.; Watson, B. S.; Sumner, L. W. Analytical and biological variances associated with proteomic studies of *Medicago truncatula* by two-dimensional polyacrylamide gel electrophoresis. *Proteomics* **2002**, *2* (8), 960–968.
- (43) Newman, J. R.; Ghaemmaghami, S.; Ihmels, J.; Breslow, D. K.; Noble, M.; DeRisi, J. L.; Weissman, J. S. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **2006**, *441* (7095), 840–846.

PR0700450