Plant Physiology and Biochemistry 215 (2024) 109032

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Comparative proteomic analyses of potato leaves from field-grown plants grown under extremely long days

Svante Resjö [a,*,#], Jakob Willforss [b,#], Annabel Large [a,1], Valentina Siino [b], Erik Alexandersson [a], Fredrik Levander [b], Erik Andreasson [a]

[a] Department of Plant Protection Biology, Swedish University of Agricultural Sciences, PO Box 190, SE-234 22, Lomma, Sweden
[b] Department of Immunotechnology, Lund University, Sweden

ARTICLE INFO

Handling Editor: Shivendra Sahi

ABSTRACT

There are limited molecular data and few biomarkers available for studies of field-grown plants, especially for plants grown during extremely long days. In this study we present quantitative proteomics data from 3 years of field trials on potato, conducted in northern and southern Sweden and analyze over 3000 proteins per year of the study and complement the proteomic analysis with metabolomic and transcriptomic analyses. Small but consistent differences linked to the longer days (an average of four more hours of light per day) in northern Sweden (20 h light/day) compared to southern Sweden can be observed, with a high correlation between the mRNA determined by RNA-seq and protein abundances. The majority of the proteins with differential abundances between northern and southern Sweden could be divided into three groups: metabolic enzymes (especially GABA metabolism), proteins involved in redox metabolism, and hydrolytic enzymes. The observed differences in metabolic enzyme abundances corresponded well with untargeted metabolite data determined by GC and LC mass-spectrometry. We also analyzed differences in protein abundance between potato varieties that performed relatively well in northern Sweden in terms of yield with those that performed relatively less well. This comparison indicates that the proteins with higher abundance in the high-yield quotient group are more anabolic in their character, whereas the proteins with lower abundance are more catabolic. Our results create a base of information about potato "field-omics" for improved understanding of physiological and molecular processes in field-grown plants, and our data indicate that the potato plant is not generally stressed by extremely long days.

1. Introduction

Proteomics and other omics data sets of field-grown crops are rare, but they represent an important base of information needed to understand molecular and systems field biology ("field-omics") (Alexandersson et al., 2014). Plants grown during extremely long days may require special properties to fully benefit from the high amount of available light, but there are very few molecular data or biomarkers available regarding the growth of field-grown plants under such conditions. For example, the growing season in agricultural areas in northern Sweden has days that are approximately 20 h long. The average temperature in Scandinavia is predicted to increase considerably. According to a recent estimate, the mean temperature during 2070–2100 will be 2–4 °C higher than during 1970–2000 under the RCP4.5 greenhouse gas concentration trajectory (Jacob et al., 2014). A consequence of this is that the borders for economically viable cultivation of a number of crops, including potato, might be shifted northward. With climate change, the relative importance of Nordic agriculture is expected to increase, which calls for more latitudinal studies measuring both the performance and molecular mechanisms at play (Roitsch et al., 2022).

Potatoes are the third most important crop globally (Food and Agriculture Organization of the United Nations), and in many food cultures potatoes are difficult to replace. The potato plant originated in the Andean highlands close to the equator, where the days are always close to 12 h in length, the temperature is relatively low (15–18 °C), and

* Corresponding author.
E-mail address: svante.resjo@slu.se (S. Resjö).
[1] Current address: San Francisco Graduate Program in Bioengineering, University of California, Berkeley, CA, USA.
[#] Svante Resjö and Jakob Willforss contributed equally to this work.

https://doi.org/10.1016/j.plaphy.2024.109032
Received 11 January 2024; Received in revised form 8 May 2024; Accepted 7 August 2024
Available online 10 August 2024
0981-9428/© 2024 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

the growing season is year round (Haverkort, 1990). The ancestors of modern potato cultivars were thus, adapted to these conditions. Potato cultivars that are better adapted to higher temperatures, longer days, and shorter growing seasons have been developed. In particular, the adaption to long days (approximately 14–16 h) made it possible to grow potatoes in temperate climates like those in Europe. While the average temperature in most of Sweden is close to the temperatures to which the ancestral potato plants were adapted, the days can be considerably longer, and the growing season is shorter, both compared to the Andean site of origin and to many other areas where potatoes are grown. The influence of latitude on potato cultivation has mainly been studied at lower latitudes (i.e., 0–50 °N). In Sweden, potatoes are mainly cultivated in the south, between 55° N and 59° N, since the climate is currently more favorable for potato cultivation at these latitudes (Eriksson et al., 2016). However, potato cultivation is possible further north, for example in Umeå at 64° N with 20 h light during a large part of the growing season, and cultivation of seed potato is done there owing to the lower pathogen pressure (Eriksson et al., 2016). A future breeding goal could be to extend the areas where crop cultivation is economically viable with genotypes that can take full advantage of the long days.

To achieve this, it would be advantageous to improve our understanding of two fundamental biological questions: Firstly, what are the molecular effects of potato cultivation during the extreme long days of summer in northern Sweden? Secondly, what are the molecular characteristics of potato cultivars that take advantage of the longer days? Increased knowledge about molecular mechanisms that are linked to higher potato yields in the northern hemisphere could be useful in future selection of cultivars, breeding, gene editing, cultivation practice, and systems biology. For these purposes, molecular analyses of field-grown materials are informative (Alexandersson et al., 2014). Studies of field apoplastic proteomes from field grown potato has focused on seasonal variation (Lankinen et al., 2018; Abreha et al., 2021) Transcriptome and metabolite markers were previously identified from 31 potato cultivars grown under optimal and reduced water supply in six independent field trials with the help of machine learning (Sprenger et al., 2018). A number of SSR markers were later constructed based on the transcript biomarker discovery and successfully employed (Schumacher et al., 2021).

By combining high-throughput techniques and integrative analyses, potato -omics approaches are providing more holistic views of the potato pangenome, gene expression patterns, protein composition, and metabolic pathways. For example, Boutsika et al. recently did a combinatory exploration of the transcriptome, proteome, methylome and microbial composition of potatoes of the same cultivar grown on the mainland Greece and Naxos, which has the status of Protected Geographical Indication. They found differences in both primary metabolism (e.g. sucrose synthases) and plant defence expression profile (Kunitz trypsin inhibitors). They also concluded that potatoes grown on Naxos recruits more microbes thought to be beneficial (Boutsika et al., 2023).

We present proteomics data from field trials on potatoes in northern and southern Sweden, respectively, from three separate years. These findings are complemented by metabolomic and transcriptomic analyses of selected field samples. We also examine the characteristics of varieties that yield well in northern Sweden compared to varieties that yield relatively less well.

2. Methods

2.1. Potato cultivation and sampling

Potatoes were grown in Borgeby in southern Sweden, at 55.75° N, 13.05° E, and in Umeå in northern Sweden, at 63.81° N, 20.23° E, in the years 2016, 2018, and 2019. Seed potatoes were purchased from certified seed potato providers, with the exception of the genotypes 913057 and 908510, which were obtained from the Swedish potato breeding program at SLU Alnarp. The cultivation of potato was done in essentially the same manner as that described in (Liljeroth et al., 2010). The development of the plants was monitored, and leaflets were sampled at matched developmental stages between the two sites. For details about sampling dates, see Supplemental Table 1. Eight leaflets were sampled from the third leaf from the top, wrapped in aluminum foil, and snap frozen in liquid nitrogen immediately after sampling. From 2016, 63 samples from 12 varieties collected in July and August were analyzed (Fig. 1). In addition, samples were obtained from field trials in 2018 for both July and August (12 samples from each month) and July 2019 (10 samples). Owing to extreme weather conditions during July 2018, this set of samples was omitted from the analysis.

2.2. Protein extraction and digestion with trypsin

Proteins were extracted from potato leaves using the phenol extraction methanol/ammonium acetate precipitation method described in (Carpentier et al., 2005) with a modified resuspension

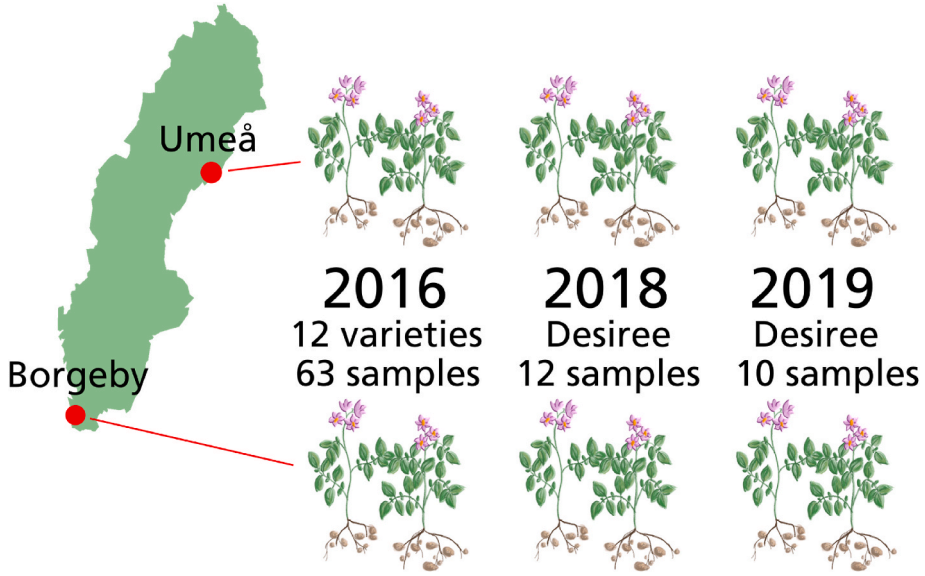

Fig. 1. Overview of the test sites and sampling strategy.

buffer. Briefly, leaves were frozen in liquid nitrogen and ground to a powder in a mortar. Fifty milligram of leaf powder was transferred to a 2 mL tube, and 500 μL of extraction buffer (50 mM Tris-HCl pH 8.5, 5 mM EDTA, 100 mM KCl, 1% DTT, 30% sucrose, 1 mM PMSF, protease inhibitor cocktail (Sigma P9599)) was added. The proteins were precipitated according to (Carpentier et al., 2005), but the final resuspension buffer consisted of 50 mM Tris-HCl pH 7,5, 2 mM EDTA, 4% SDS, 1 mM DTT, 1 mM PMSF, Protease inhibitor cocktail (Sigma P9599). The samples were then stored at −80 °C. For tryptic digestion, DTT was added to the sample to a final concentration of 10 mM, followed by a 10 min incubation at 95 °C. Subsequently, iodoacetamide was added to a final concentration of 55 mM, and the samples were incubated for 30 min in the dark. The samples were then analyzed by SDS-PAGE until the front had migrated approximately 5 mm into the gel. The pieces were excised, washed, and the protein was digested with trypsin added to the gel. The tryptic peptides were cleaned with C18 columns (UltraMicro spin columns, Nest Group Inc., Southborough, MA, USA) and resuspended in a solution of 5% formic acid in water.

2.3. Proteome mass spectrometry (MS) analysis

Peptide digests from the 2018 and 2019 samples were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) using an EASY-nano LC system (Thermo Fisher Scientific, Germany) coupled with a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific, Germany) operating in positive ion mode for data-dependent acquisition. The analytical column was a 15 cm long fused silica capillary (75 μm × 16 cm Pico Tip Emitter, New Objective), packed in house with C18 material ReproSil-Pur 1.9 μm (Dr. Maisch GmbH, Germany). Peptides were separated by an 80 min gradient from 5% to 90% solvent B (80% ACN, 0.1% FA) at a constant flow rate of 250 nL/min. The Orbitrap acquired the full MS scan with an automatic gain control target value of 3 × 106 ions and a maximum fill time of 50 ms. The 20 most abundant peptide ions were selected from the MS for higher energy collision-induced dissociation fragmentation (collision energy: 40 V). Fragmentation was performed at 15,000 FWHM resolution for a target of 1 × 105 and a maximum injection time of 20 ms using an isolation window of 1.2 m/z. Xcalibur software v 3.0 (Thermo Fisher Scientific, Germany) was used to control the nLC system and the MS and to acquire and visualize the raw data.

For the 2016 samples, prior to MS analyses, 80 μL of defrosted peptide extracts were acidified with 5 μL of 10% formic acid and then trapped and enriched on StageTip columns (Rappsilber et al., 2007). Next, peptides were extracted with 70% ACN and 0.1% TFA and dried down before reconstitution in 20 μL of 2% ACN and 0.1% TFA. MS analyses were carried out on an Orbitrap Fusion Tribrid MS system (Thermo Scientific) equipped with a Proxeon Easy-nLC 1000 (Thermo Fisher). Injected peptides were trapped on an Acclaim PepMap C18 column (3 μm particle size, 75 μm inner diameter × 20 mm length). After trapping, gradient elution of peptides was performed on an Acclaim PepMap C18 column (100 Å 3 μm, 150 mm, 75 μm). The outlet of the analytical column was coupled directly to the mass spectrometer using a Proxeon nanospray source. The mobile phases for LC separation were 0.1% (v/v) formic acid in LC-MS grade water (solvent A) and 0.1% (v/v) formic acid in acetonitrile (solvent B). Peptides were first loaded onto the trapping column using solvent A and a constant pressure mode. Subsequently, peptides were eluted via the analytical column at a constant flow rate of 300 nL/min. During the elution steps, the percentage of solvent B increased from 5% to 10% in the first 2 min, then increased to 25% in 50 min, then to 60% in 15 min, and to 90% in a further 5 min, where it was maintained for 5 min. The peptides were introduced into the mass spectrometer via a stainless steel emitter 40 mm (Thermo Fisher), and a spray voltage of 1.9 kV was applied. The capillary temperature was set to 275 °C. Data acquisition was carried out using a top N based data-dependent method with a cycle time of 3 s. The master scan was performed in the Orbitrap in the range of 350–1350 m/z at a resolution of 120,000 FWHM. The filling time was set at a maximum of 50 ms with a limitation of 4 × 105 ions. Ion trap CID-MS2 was acquired using parallel mode, with a maximum filling time of 300 ms with a limitation of 2 × 103 ions, a precursor ion isolation width of 1.6 m/z, and a resolution of 15,000 FWHM. The normalized collision energy was set to 35%. Only multiply charged (2+ to 5+) precursor ions were selected for MS2. The dynamic exclusion list was set to 30 s and a relative mass window of 5 ppm. After inspection of the principal component analysis, a batch effect related to the switch of the liquid chromatography column in the HPLC system was identified, leading to the rerun of 33 samples; this was done to avoid statistical comparisons for conditions that, prior to reruns, were unbalanced between the batches.

2.4. Processing of proteomics data

Raw MS data in Thermo raw format were converted to mzML and MGF using Proteowizard (Chambers et al., 2012) version 3.0.11841 (Q Exactive data) or 3.0.9220 (Fusion data) with MS-Numpress compression (Teleman et al., 2014), and MS1 peptide features were detected using Dinosaur (Teleman et al., 2016) version 1.1.3. Processing of mass spectra to peptide abundances was performed in the Proteios software environment (Hakkinen et al., 2009) 2.20.0-dev, using builds 4646 and 4626 for the two datasets, respectively. MS/MS peptide fragment fingerprinting was performed using MS-GF+ (Kim and Pevzner, 2014) and X!Tandem Alanine (2017.2.1.4, http://www.thegpm.org/TANDEM/)) against a database consisting of the potato proteome available in UniProt as of January 17, 2017, including the canonical sequences and isoforms, that was expanded with an equal number of decoy (reverse sequence) protein entries; the parameters for MS/MS peptide fragment fingerprinting were 10 ppm precursor tolerance, fixed carbamidomethylation of C, variable oxidation of M, and protein N-terminal acetylation. The “instrument” search parameter was set to Q Exactive for Q Exactive data and LowRes in MS-GF + for Fusion data, whereas the corresponding fragment settings in X!Tandem were 0.02 Da and 0.4 Da, respectively. The search results were combined in Proteios, and peptides passing a peptide-spectrum-match (PSM) q-value threshold of 0.001 were matched against the MS1 features with an m/z tolerance of 0.004, also including PSMs with q-values <0.1 if the same peptide was identified elsewhere in the dataset at q < 0.001. Features were aligned, and peptide feature identities were propagated between files in Proteios using default settings before export of the peptide feature table for further analysis. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Perez-Riverol et al., 2022) with the dataset identifiers PXD026661 and PXD026690.

2.5. Outlier removal

The data were inspected for outliers and trends using the interactive software OmicLoupe (v0.9.7) (Willforss et al., 2021). In the 2016 dataset, two strong outliers were identified using density plots. These samples were omitted from subsequent analyses. In the July 2018 dataset, one sample was identified as an outlier using a dendrogram, and for 2019 data, one sample was identified as a strong outlier based on having a strongly different intensity profile as seen in a density plot.

2.6. Proteome abundance matrix processing and statistics

Raw peptide intensities were normalized using cyclic Loess normalization (Ballman et al., 2004) in NormalyzerDE 1.5.4, (Willforss et al., 2019). Peptides eluted during the post-gradient column wash were removed prior to further data analysis (i.e., those with retention times >63 min in the 2016 dataset and >69.5 min in the combined 2018 and 2019 dataset). Protein intensities were subsequently calculated from peptide intensities using an R implementation (v0.9.3, https://github.co

m/ComputationalProteomics/ProteinRollup) of the RRollup algorithm in DanteR (Polpitiya et al., 2008) version v1.1.6970. Differential expression between groups was calculated using Limma v3.42.2 (Ritchie et al., 2015), a statistical approach in which variance from the whole dataset is used to improve variance estimates for individual proteins, to calculate the contrasts between the Umeå and Borgeby field sites for the potato varieties Desiree, Rocket, and Bintje in 2016, using whether the samples had been processed before or after the column swap in the mass spectrometer as a covariate. Here, the final mass spectrometry rerun for each sample was used, which best balanced samples across the column swap. Furthermore, comparisons were made between Umeå and Borgeby for Desiree in 2018 and 2019. Proteins with false discovery rate (FDR) < 0.1 across all three years were selected as stable candidates. In order to study the relationship between yield and expression, varieties with high (above 1) and low (below 0.75) yield quotients for Umeå/-Borgeby were identified and selected, resulting in the varieties Amour, 908510, 913057, Sarpomira, Mandel, and Folva being classified as 'HIGH' and Carolus, Solist, and Artemis being classified as 'LOW'. These groups are outlined in Supplemental Table 2. A statistical comparison between the HIGH and LOW groups was performed using Limma, again using the column swap as a covariate. In this case, the original MS run for each sample was used, as, in this case, this best balanced the samples across the column swap. Proteins with FDR <0.05 were considered significant. The proteins were further annotated by performing homology pBLAST searches (2.6.0+, 10.1016/S0022-2836(05)80360-2) against the *Solanum tuberosum* reference genome protein sequences (version 4.03). In each case, the annotation was transferred from the best homology match.

2.7. RNA extraction

Leaflets were flash frozen in liquid nitrogen and then homogenized to a fine powder using a mortar and pestle. RNA extractions were performed using an RNeasy Plant Mini kit (Qiagen GmbH, Hilden, Germany) and approximately 100 mg of leaf tissue. Samples were treated with DNase. RNA concentration and purity (260/280 nm > 1.8) were checked using an ND-1000 NanoDrop (Wilmington, USA), and the integrity of the samples was analyzed with an Experion Automated Electrophoresis System (Bio-Rad Laboratories, Hercules, USA).

2.8. RNA-seq data

RNA sequencing (RNA-seq) was performed using the SNP & SEQ Technology Platform in Uppsala, Sweden with NovaSeq S4 flow cells, paired-end 150 bp read lengths, and v1 sequencing chemistry. Sequencing libraries were prepared from 750 ng total RNA using the TruSeq stranded total RNA library preparation kit with RiboZero Gold treatment (cat# 20020598/9, Illumina Inc.). Library preparation was performed according to the manufacturer's protocol (# 1000000040499). Trimmomatic (Bolger et al., 2014) was used for the quality trimming of reads (version 0.36), trimming TruSeq3 adapters with settings "TruSeq3-PE-2.fa:2:30:10" and with quality sliding window settings "SLIDINGWINDOW:5:20", while filtering reads shorter than 50 nucleotides. SortmeRNA (Kopylova et al., 2012) (version 2.1b, was used against bacterial and eukaroytic ribosomal RNA databases with default settings to filter ribosomal RNAs. Bowtie2 (version 2.4.1) (Langmead and Salzberg, 2012) was used to filter non-coding RNA using the plant noncoding RNA database (http://structuralbiology.cau.edu.cn/PNRD/download.php) under default settings except for running with the local alignment setting. The reads were quality controlled using FastQC (v0.11.9, https://qubeshub.org/resources/fastqc) and MultiQC v1.8 (Ewels et al., 2016) for combined quality reports. They were subsequently mapped using STAR v2.7.3 (Dobin et al., 2013) to the PGSC potato reference genome v4.03 (Hardigan et al., 2016). For the potato reference genome, the annotation was downloaded as a GFF file and subsequently parsed into GTF format using gffread in Cufflinks version 2.2.1 (Trapnell et al., 2012). STAR was run using quantMode GeneCounts based on the GTF file. The resulting count data were combined into a count matrix and used for subsequent analysis. DESeq2 version 1.26.0 (Love et al., 2014) was run using default settings and was used to calculate contrasts between Umeå and Borgeby for the varieties Desiree, Bintje and Rocket. The final output was further inspected for outliers and trends using the interactive visualizations provided by OmicLoupe, with no outliers removed. The RNAseq data has been deposited to the SRA database at the NCBI biobank with the accession number PRJNA1002643 and can be accessed at https://www.ncbi.nlm.nih.gov/sra/PRJNA1002643.

2.9. Metabolomics analysis

Metabolic profiling by gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS) was performed at the Swedish Metabolomics Center in Umeå, Sweden. Information about reagents, solvents, standards, reference and tuning standards, and stable isotope internal standards can be found in Supplementary Materials 6.

Sample Preparation: Sample preparation was performed according to (Gullberg et al., 2004). Leaves were frozen in liquid nitrogen and ground to a powder in a mortar, and 20 mg (±10%) of frozen leaf powder was used. The samples were analyzed in batches (different sample types) according to a randomized run order on both GC-MS and LC-MS.

GCMS Analysis: Derivatization and GCMS analyses were performed as described previously (Gullberg et al., 2004). An aliquot of 0.5 μL of each derivatized sample was injected in splitless mode by an L-PAL3 autosampler (CTC Analytics AG, Switzerland) into an Agilent 7890B gas chromatograph equipped with a 10 m × 0.18 mm fused silica capillary column with a chemically bonded 0.18 μm Rxi-5 Sil MS stationary phase (Restek Corporation, U.S.) The injector temperature was 270 °C. The purge flow rate was 20 mL min^{-1}, and the purge was turned on after 60 s. The gas flow rate through the column was 1 mL min^{-1}; the column temperature was held at 70 °C for 2 min, then increased by 40 °C min^{-1} to 320 °C, and held there for 2 min. The column effluent was introduced into the ion source of a Pegasus BT time-of-flight mass spectrometer (GC/TOFMS, Leco Corp., St Joseph, MI, USA). The transfer line and ion source temperatures were 250 °C and 200 °C, respectively. Ions were generated by a 70 eV electron beam at an ionization current of 2.0 mA, and 30 spectra s^{-1} were recorded in the mass range m/z 50–800. The acceleration voltage was turned on after a solvent delay of 150 s. The detector voltage was 1800–2300 V.

LC-MS analysis: Before LCMS analysis, the sample was re-suspended in 10 μL + 10 μL methanol and water. All samples were first analyzed in positive mode. Thereafter, the instrument was switched to the negative mode, and a second injection of each sample was performed. Chromatographic separation was performed on an Agilent 1290 Infinity UHPLC-system (Agilent Technologies, Waldbronn, Germany). Two microliters of each sample were injected onto an Acquity UPLC HSS T3, 2.1 × 50 mm, 1.8 μm C18 column in combination with a 2.1 mm × 5 mm, 1.8 μm VanGuard precolumn (Waters Corporation, Milford, MA, USA) held at 40 °C. The gradient elution buffers were A (H_2O, 0.1 % formic acid) and B (75/25 acetonitrile:2-propanol, 0.1 % formic acid), and the flow rate was 0.5 mL min^{-1}. The compounds were eluted with a linear gradient consisting of 0.1–10 % B over 2 min, B was increased to 99 % over 5 min and held at 99 % for 2 min; B was decreased to 0.1 % for 0.3 min, and the flow rate was increased to 0.8 mL min^{-1} for 0.5 min; these conditions were held for 0.9 min, after which the flow rate was reduced to 0.5 mL min^{-1} for 0.1 min before the next injection.

The compounds were detected with an Agilent 6550 Q-TOF mass spectrometer equipped with a jet stream electrospray ion source operating in positive or negative ion mode. The settings were identically maintained across both modes, with the exception of the capillary voltage. A reference interface was connected for accurate mass measurements. The reference ions purine (4 μM) and HP-0921 (Hexakis(1H,

1H, 3H-tetrafluoropropoxy)phosphazine) (1 μM) were infused directly into the MS at a flow rate of 0.05 mL min^{-1} for internal calibration, and the monitored ions were purine m/z 121.05 and m/z 119.03632; HP-0921 m/z 922.0098 and m/z 966.000725 for the positive and negative modes, respectively. The gas temperature was set to 150 °C, the drying gas flow was set to 16 L min^{-1}, and the nebulizer pressure was 35 psig. The sheath gas temperature was set to 350 °C and the sheath gas flow was 11 L min^{-1}. The capillary voltage was set to 4000 V in positive ion mode and to 4000 V in negative ion mode. The nozzle voltage was 300 V. The fragmentor voltage was 380 V, the skimmer was 45 V, and the OCT 1 RF Vpp was 750 V. The collision energy was set to 0 V. The m/z range was 70–1700, and data were collected in centroid mode with an acquisition rate of 4 scans s^{-1} (1977 transients/spectrum).

2.10. Handling of metabolomics data

For the GC-MS data, all non-processed MS files from the metabolic analysis were exported from the ChromaTOF software in NetCDF format to MATLAB R2016a (Mathworks, Natick, MA, USA), where all data pretreatment procedures, such as baseline correction, chromatogram alignment, data compression, and multivariate curve resolution, were performed using custom scripts. The extracted mass spectra were identified by comparing their retention index values and mass spectra with libraries of retention time indices and mass spectra (Schauer et al., 2005). Mass spectra and retention index comparisons were performed using NIST MS 2.0 software. The annotation of mass spectra was based on reverse and forward searches in the library. Masses and the ratio between masses indicative of a derivatized metabolite were especially noted. If the mass spectrum, according to the SMC experience, had the highest probability indicative of a metabolite, and the retention index between the sample and library for the suggested metabolite was ±5 (usually less than 3), the deconvoluted "peak" was annotated as an identification of a metabolite.

For the LC-MS data, all data processing was performed using the Agilent Masshunter Profinder version B.08.00 (Agilent Technologies Inc., Santa Clara, CA, USA). Processing was performed in both targeted and untargeted fashions. For target processing, a predefined list of metabolites (including amino acids, bile acids, acyl-carnithines, fatty acids, lyso-phosphatidylcholines, nucleotides, short-length peptides, and steroids, among others) commonly found in plasma and serum were searched for using the batch targeted feature extraction in Masshunter Profinder. An in-house LC-MS library (Gullberg et al., 2004), assembled using authentic standards run on the same system with the same chromatographic and mass-spectrometry settings, was used for the targeted processing. The identification of the metabolites was based on MS, MSMS, and retention time information. For the untargeted data, each tissue group was processed individually using the batch recursive feature extraction algorithm within Masshunter Profinder.

The metabolomics data were further processed using NormalyzerDE (Willforss et al., 2019). Based on the normalization quality report provided by NormalyzerDE, cyclic Loess normalization (Ballman et al., 2004) was used as the normalization method. Upon inspection of density plots in OmicLoupe (Willforss et al., 2021), one sample was identified as a strong outlier and omitted from the analysis. Subsequently, differential expression was calculated using Limma in NormalyzerDE, comparing Desiree samples between Umeå and Borgeby, comparing all varieties between Umeå and Borgeby, and comparing the HIGH versus LOW yield groups within Umeå and Borgeby (grouped as in the proteomics analysis, outlined in Supplementary Material 2).

All R analyses used to perform the proteomic and RNA-seq analyses, and the code used to generate figures for this manuscript are presented as an R Markdown document in Supplementary Material 3.

3. Results and discussion

3.1. Potato leaf data from northern and southern Sweden

We have generated molecular data from 3 years (2016, 2018, and 2019) of field trials in the north and south of Sweden (Fig. 1). This manuscript is focused on the analysis of proteins that consistently differ in abundance between leaves from potatoes of the cultivar Desiree grown at the southern and northern sites over the entire period (Fig. 2). The analysis of the 2016 dataset resulted in the identification of 30,077 peptides and 3402 proteins, while the combined analysis of the 2018 and 2019 datasets resulted in the identification of 24,336 peptides and 3598 proteins. The differentially abundant proteins that passed FDR <0.1 in all 3 years are listed in Table 1 and Supplementary Material 4.

Twenty-two proteins were found to be both significantly and consistently differentially abundant between the northern (64° N) and southern (56° N) field sites. The abundance of the corresponding transcripts and proteins was also analyzed for the varieties Desiree, Bintje and Rocket from 2016 to further validate the multi-seasonal proteomics data (Fig. 3). The protein abundance of the corresponding proteins in Bintje and Rocket correlated very well with abundance in Desiree, and in all cases in which the transcript abundance was significantly different, it varied in the same direction as the protein abundance (Fig. 3). This correlation between transcript and protein abundance across cultivars and growing seasons strengthens our conclusion that these differences reflect robust effects of the growth location. It is also in line with the high level of correspondence between the transcriptome and proteome datasets identified by Boutsika et al. (2023) in field grown potatoes where all pairwise comparisons of both transcriptome and proteome were positive in one cultivar between the two sites they studied. With the exception of this study, there are no other -omics studies of field grown potatoes that we are aware of.

The majority of differentially abundant proteins can be divided into three groups: enzymes involved in the metabolism of amino acids and citrate, proteins involved in redox metabolism, and hydrolytic enzymes. A number of metabolites involved in primary and amino acid metabolism have higher abundance in the northern site. In addition, metabolites associated with oxidative and biotic stress have lower abundance in the north.

3.2. Citrate and amino acid/glutamate metabolism

Seven metabolic enzymes were identified as having significantly different abundances (Fig. 4). Of these, two gamma-aminobutyrate (GABA) transaminase isoform 2 (M1C906, M1CD67), alanine aminotransferase (M1A1E3), 2-oxoglutarate-dependent dioxygenase (M1AUM2), and methylthioribose kinase (M1CFQ4) are directly involved in amino acid metabolism. Of the remaining two, dihydrolipoyl dehydrogenase (M0ZSL5) is part of the pyruvate dehydrogenase complex, whereas pyruvate decarboxylase (M1D0M1) catalyzes the decarboxylation of pyruvate to acetaldehyde. Some of these enzymes are thus important for controlling the metabolic fate of pyruvate, which is a key amino acid precursor, while others catalyze transamination reactions that are a part of amino acid metabolism.

In the northern site there was a higher abundance of glutamate, succinate, alpha ketoglutarate and citrate (Fig. 4). Glutamate plays an important part in amino acid metabolism, both as a donor of amino groups and of carbon atoms for the carbon skeletons of some amino acids. A number of other amino acids also had significantly different abundances between the sites.

Glutamate is a substrate of GABA transaminase, but this enzyme accepts multiple substrates, and several of the metabolites involved can be used as substrates by more than one of the four metabolic enzymes, which makes it challenging to elucidate exactly how their regulation is interrelated. GABA transaminase catalyzes the conversion of GABA to succinate semialdehyde as a part of the GABA shunt (Michaeli and

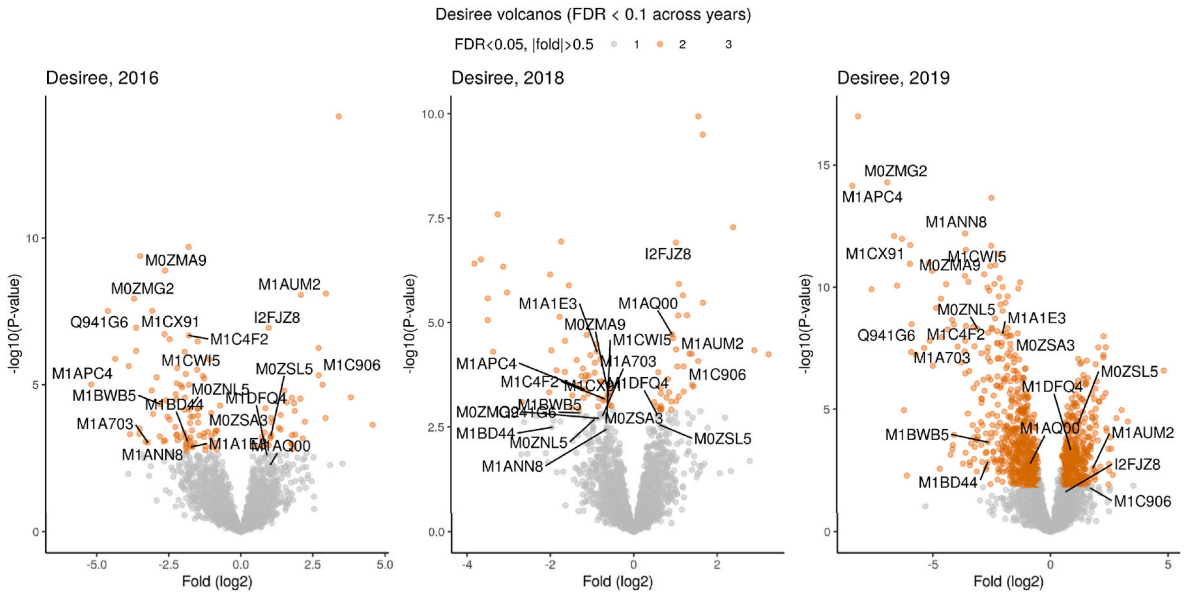

Fig. 2. Volcano plots showing Desiree proteins with FDR <0.1 across three years. Proteins found with FDR <0.1 across all years are labelled in all three plots.

Table 1
The 22 proteins found to consistently be significantly different between 64 °N and 56 °N over three seasons.

UniProt ID	Protein name	Location with highest abundance
M1APC4,M1APC7, M1APC8, M1APC9	Acidic class II 1,3-beta-glucanase	South
M1A703	NtPRp27	South
Q941G6	Cytoplasmic small heat shock protein class I	South
M0ZMG2	Class II chitinase	South
M1AY17	Cationic peroxidase	South
M1ANN8	Glutathione transferase	South
M1CX91	Glucan endo-1,3-beta-D-glucosidase	South
M0ZMA9	PR10	South
M1D0M1	Pyruvate decarboxylase	South
M1C4F2,M1C4F3	41 kD chloroplast nucleoid DNA binding protein (CND41)	South
M0ZNL5	Glutathione s-transferase	South
M1CWI5	Glutathione reductase	South
M1A1E3	Alanine aminotransferase	South
M0ZSA3	Monodehydroascorbate reductase	South
M1CFQ4	Methylthioribose kinase	South
M1DFQ4	Aspartic proteinase nepenthesin-1	North
I2FJZ8	Carbonic anhydrase	North
M0ZSL5	Dihydrolipoyl dehydrogenase	North
M1AQ00	Superoxide dismutase [Cu–Zn] 2	North
M1CD67	Gamma aminobutyrate transaminase isoform2	North
M1AUM2	2-oxoglutarate-dependent dioxygenase	North
M1C906	Gamma aminobutyrate transaminase isoform2	North

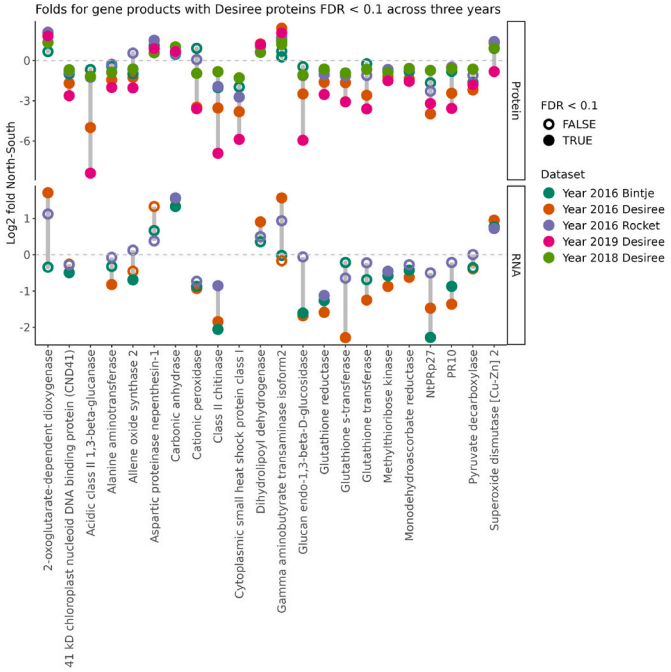

Fig. 3. Fold change of Desiree proteins with FDR <0.1 across three years, together with the corresponding proteins in the cultivars Bintje and Rocket 2016, as well as the corresponding mRNA levels in Desiree, Bintje, and Rocket in 2016. An FDR smaller than 0.1 is indicated by a filled circle while an FDR larger than 0.1 is indicated with an open circle.

Fromm, 2015; Ramos-Ruiz et al., 2019). This enzyme can use either pyruvate or alpha ketoglutarate as amine acceptors, producing alanine or glutamate, respectively.

The higher glutamate abundance in the metabolomes of samples from northern Sweden indicates that the reaction producing glutamate is the predominant one catalyzed by the increased GABA transaminase observed at that location. We also observed increases in succinate and alpha ketoglutarate in the northern site. The increased abundance of alpha ketoglutarate might be partially explained by the reduced amounts of alanine transaminase. Alanine transaminase catalyzes the reversible conversion of alanine and alpha ketoglutarate into glutamate and pyruvate. A decrease in alanine transaminase activity could contribute to the accumulation of alanine and alpha ketoglutarate. The increased alpha ketoglutarate could then be utilized by 2-oxoglutarate-dependent dioxygenase, an enzyme that uses alpha ketoglutarate and oxygen to oxidize a wide variety of substrates, producing succinate and oxidized substrates. Succinate is also produced from succinate semi-aldehyde by two successive enzymatic reactions. Thus, the differences observed in the abundances of metabolic enzymes corresponded well with the observed differences in the metabolites (Fig. 5). Succinate and

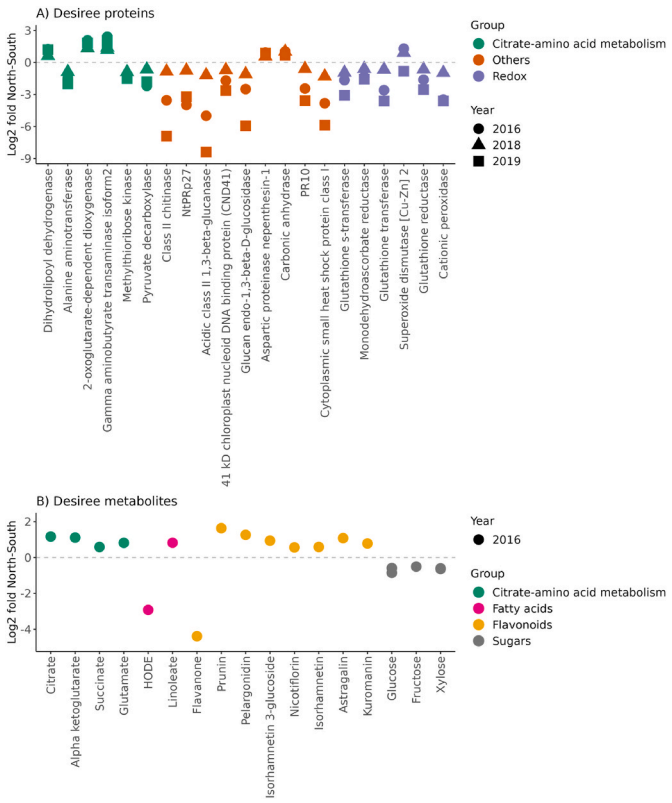

Fig. 4. Selected significant proteins and related metabolites that are discussed in the text.

alpha ketoglutarate are intermediates in the citric acid cycle. In addition, citrate was elevated in the samples from northern Sweden, indicating higher citric acid cycle activity in these plants. The increased abundance of dihydrolipoyl dehydrogenase indicates that the pyruvate dehydrogenase complex was more active in plants from northern Sweden, which would also have contributed to the increased abundance of citric acid cycle intermediates through the conversion of pyruvate to acetyl CoA by the pyruvate dehydrogenase complex. The citric acid cycle is a central component of the catabolic machinery of the cell. Increases in proteins involved in catabolism has been observed in proteomics analyses of short-term cold response in a number of plant species (Kosova et al., 2011, 2013) such as wheat and *Arabidopsis thaliana*. However, in potato short-term cold treatment has mixed results on metabolic enzymes, with some (such as ATP synthase) decreasing and others (such as glutamine synthetase) increasing (Li et al., 2021). The plants grown in northern Sweden in this study have been exposed to lower temperature on average, and it is possible that this contributes to the increased citrate and amino acid/glutamate metabolism that our data indicates. However, it should be noted that the temperature difference is much larger in the cold response studies (typically 16–18°) and the time of the study is shorter (typically one or a few weeks) than in our study. Thus, it is also possible that other differences, such as the extremely long days in northern Sweden are causing the observed effects.

3.3. Redox regulation

The next major group of proteins that consistently differed in abundance between northern and southern Sweden are six proteins involved in redox processes. Only one (superoxide dismutase [Cu–Zn] 2, M1AQ00) was more abundant in the samples from northern Sweden. Superoxide dismutase is important for protecting plants against reactive oxygen species (ROS). The other five redox related proteins, glutathione reductase, (M1CWI5), glutathione *S*-transferase (M0ZNL5), glutathione transferase (M1ANN8), monodehydroascorbate reductase (M0ZSA3), and cationic peroxidase (M1AY17), were more abundant in the south (Fig. 4). This type of ROS-scavenging enzymes have been found to be upregulated in response to cold in plants other than potato (Kosova et al., 2011), but not in potato (Li et al., 2021). These enzymes could also be upregulated in response to many different stresses.

The first three proteins are mainly involved in the detoxification of xenobiotics by conjugation with glutathione (Gullner et al., 2018). Recently, site-specific hypermethylation of glutathione S-transferase in field grown potato has been shown to affect the abundance of both protein and transcript levels in tubers grown at Naxos (Boutsika et al., 2023).

Monodehydroascorbate reductase is part of the glutathione-ascorbate cycle, which is a hydrogen peroxide detoxification pathway, and cationic peroxidase is also involved in hydrogen peroxide removal. The metabolomics data also indicate different levels of oxidative stress at the two sites (Fig. 4). Linoleate was more abundant in northern Sweden, whereas hydroxyoctadecadienoic acid was more abundant in southern Sweden. Linoleate is a polyunsaturated fatty acid that is oxidized to different hydroperoxyoctadecadienoic acids (HODE) by the free radicals produced during oxidative stress. Hydroperoxyoctadecadienoic acids are then further oxidized to hydroxyoctadecadienoic acid by glutathione peroxidases and phospholipases (Yoshida et al., 2013). This difference indicates a higher degree of lipid peroxidation at the site in southern Sweden. Another difference between the sites was the higher abundance of flavonoids and their glycosides in

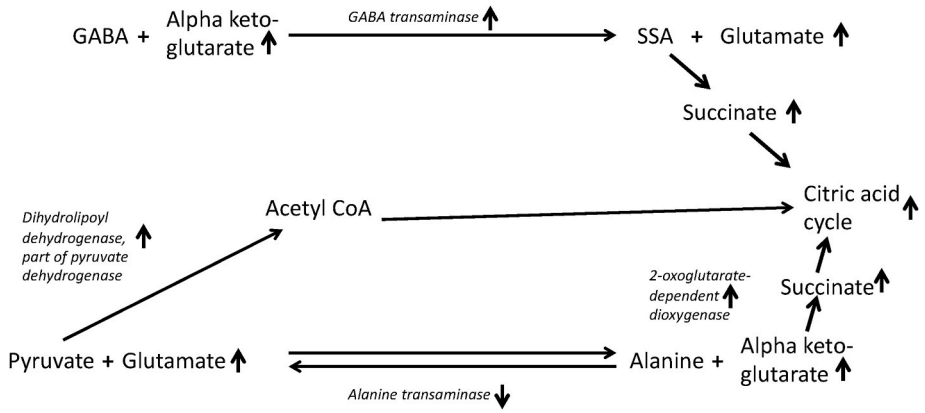

Fig. 5. The four metabolic enzymes consistently displaying differential abundance between the northern and southern site and their associated metabolites. An arrow pointing upward means that the relative abundance is higher in northern Sweden, whereas an arrow pointing downward means that the relative abundance is lower in northern Sweden.

plants from northern Sweden. Seven different flavonoid glycosides were more abundant in these samples. It is likely that this is caused by the increased length of the photoperiod at the higher latitude, since growth at high latitudes has been shown to cause increased flavonoid content in a number of plant tissues (Jaakola and Hohtola, 2010; Molmann et al., 2021).

3.4. Stress response and hydrolysis

A number of proteins involved in that can be induced by different stressors or are involved in hydrolysis were among the differentially abundant proteins (Fig. 4). Some of these are defense-related proteins that can be induced by different stressors: cytoplasmic small heat shock protein class I (Q941G6), PR10 (M0ZMA9), and NtPRp27 (M1A703). Small heat shock protein class I has been shown to increase heat stress tolerance (McLoughlin et al., 2016), while NtPRp27 has been shown to contribute to resistance against *Phytophthora infestans* (Shi et al., 2012). These proteins consistently had higher abundance in the samples from southern Sweden,

Among the proteolytic enzymes, only one, aspartic proteinase nepenthesin-1 (M1DFQ4), was more abundant in the samples from northern Sweden than those from southern Sweden. All the other enzymes exhibited the opposite pattern, i.e., higher abundance in southern Sweden (Fig. 4). One of the proteins in this group, M1C4F2, a 41 kD chloroplast nucleoid DNA binding protein (CND41), is also an aspartic protease that is involved in senescence (Kato et al., 2004). The other three are chitinases or glucanases that are likely responsible for cell wall degradation (acidic class II 1,3-beta-glucanase (M1APC4), class II chitinase (M0ZMG2), and glucan endo-1,3-beta-D-glucosidase (M1CX91). This data may reflect differences in various biotic or abiotic stresses. Increased carbohydrates have been indicated as positive regulators of defense-related genes (Rojas et al., 2014), and glucose levels were found to be higher in southern Sweden than in northern Sweden (Supplementary Material 5).

3.5. Potato leaf data comparing yield groups

In order to compare potato cultivars that perform relatively well in northern Sweden with cultivars that perform less well, we grouped the cultivars according to their yield quotients for North/South. The high-yielding cultivars had quotients above 1 (and thus perform relatively well in northern Sweden) while the low-yielding groups had quotients below 0.75 (and perform relatively less well in northern Sweden). The genotypes are described in Supplementary Material 2.

A total of 56 proteins had significantly different abundances between the groups (Table 2), with generally relatively small differences. Of these, 40 were found to have higher abundance among the varieties in the high-yield quotient group, while 16 were found to have lower abundance in the high-yield group (Fig. 6). The protein profiles of the two groups were strikingly different. Among the 40 proteins that were more abundant in the high-yield quotient group, seven were either chaperones or otherwise involved in protein folding (e.g., protein disulfide-isomerase (M1AZ99) and peptidyl-prolyl cis-trans isomerase (Q9XF12)). Another four were ribosomal proteins. Another, more diverse group can be said to be involved in the biosynthesis of complex molecules. UDP-glucose 6-dehydrogenase (M0ZM42) is part of nucleotide-sugar biosynthesis, whereas isopentenyl-diphosphate delta-isomerase (M1AB35) participates in chlorophyll biosynthesis, and diphosphomevalonate decarboxylase (M1C7S0) is involved in the production of polyisoprenoids and sterols from acetyl-CoA. This means that more than a third of the differentially abundant proteins with high abundances in the high-yield quotient group are biosynthetic enzymes. We have previously correlated yield with transcriptomes of a progeny potato population, and found a negative correlation with an Acyl-CoA synthetase over two years field studies (Alexandersson et al., 2020).

Among the 16 proteins with lower abundance in the high-yielding

Table 2
The 56 proteins found to differ between the high and low yield groups (log2, positive values indicate higher abundance in the high yield group).

UniProt ID	Protein name	Abundance in the high yield group
M1B668,M1B670,M1BWS6, M1D2J4	Glutathione S-transferase/ peroxidase	3.14
M0ZVA9,M0ZVB0,M1D5H9	Oxidoreductase, 2OG-Fe(II) oxygenase family protein	3.07
M1BWB5	Hsr203J	1.95
M1CU69,M1CU70	Mta/sah nucleosidase	1.84
M1B668,M1D2J3,M1D2J4	Glutathione S-transferase/ peroxidase	1.81
M0ZTN0,M0ZTN1	Osmotin	1.75
M0ZMA8,M0ZMA9	TSI-1 protein	1.75
M1B2S2	Cytochrome *b*6-f complex iron-sulfur subunit, chloroplastic	1.60
M0ZRT6,M1A7X7	Ribosomal protein L28	1.55
M0ZRQ9	Nonclathrin coat protein zeta1-COP	1.45
M0ZKX2	Allene oxide synthase 2	1.42
M1BIX8	Reticuline oxidase	1.40
M1DWM0	Conserved gene of unknown function	1.27
M0ZX98,M0ZXE4,M1AXK1, M1AXK2,M1AXK3, M1AXP1,M1AXP2, M1AXR0,M1BBG7, M1DCI3,M1DUG5, M1DWT5	P69F protein	1.21
M1AEG5	Glucosyltransferase	1.12
M1AZ99	Protein disulfide isomerase L-2	1.08
M0ZLG7,M0ZYR0,M1A1P3, M1AGX9,M1ATR5	Alpha-tubulin	1.06
B5M4B1,P46263,P46264	Beta-tubulin	1.04
M1AUM0	Cytochrome P-450	1.04
M1B582,M1B583,M1C075, M1CX21	Heat shock cognate protein 80	0.98
M0ZLG7,M0ZYR0,M1A1P4, M1AGX9,M1ATR5, M1D0L3,M1D0L4	Alpha-tubulin	0.91
M0ZLG7,M0ZYR0,M1A1P3, M1AGX9,M1ATR5, M1D0L3,M1D0L4, M1D8T0	Alpha-tubulin	0.91
M0ZM42,M1CCS3	UDP-glucose dehydrogenase 2	0.89
M0ZLG7,M1A1P4,M1AGX9, M1ATR5	Alpha-tubulin	0.86
M1BSN9,M1BSP0	Translocon-associated protein beta family protein	0.85
M1AP63,M1AVI1, M1BQM8,M1BY26, M1BY27	Transitional endoplasmic reticulum ATPase	0.84
M1C7S0	Mevalonate disphosphate decarboxylase	0.83
M1D610	Proteasome subunit beta type-3-A	0.80
M1CX21	Molecular chaperone Hsp90-1	0.79
M0ZM42	UDP-glucose dehydrogenase 2	0.76
M1AWJ2,M1BLB0,M1BQB7	Luminal-binding protein	0.72
M1BB68	40S ribosomal protein S19	0.71
M1BC65	Anthranilate N-benzoyltransferase protein	0.67
M1CQQ1	Peroxiredoxin	0.66
M1AB35,M1C547	Plastid isopentenyl diphosphate isomerase	0.63
Q2V999	40S ribosomal protein S15	0.55
M1BRF8	Thaliana 60S ribosomal protein L7	0.53
Q9XF12	Peptidyl-prolyl cis-trans isomerase	0.49
M1B4H0	Ebna2 binding protein P100	0.45
M1BQI2	Heat shock protein 90	0.34
M1CL86	Dihydrolipoyl dehydrogenase	−0.40
M1BPR5	Ribose-5-phosphate isomerase	−0.43
M0ZQW3,M0ZT87	Chaperonin 21	−0.59

(*continued on next page*)

Table 2 (*continued*)

UniProt ID	Protein name	Abundance in the high yield group
M0ZU27,M0ZU28,M1AVJ2	Fructose-bisphosphate aldolase	−0.74
M1ACZ6,P04045	Alpha-1,4 glucan phosphorylase L-1 isozyme, chloroplastic/amyloplastic	−0.74
M1C3R6	Uncharacterized aarF domain-containing protein kinase, chloroplastic	−0.77
P04045	Alpha-1,4 glucan phosphorylase L-2 isozyme, chloroplastic/amyloplastic	−0.79
M1BM79	Ammonium transporter 1 member 3	−0.82
M1A521,M1A522,M1A524	Conserved gene of unknown function	−0.83
M1BT07	Sua5	−0.86
M1BGQ5,Q43839	Glucose-6-phosphate 1-dehydrogenase, chloroplastic	−0.92
M1B1I3	GDSL-lipase protein	−1.06
M1CQY0,M1CQY1	Alpha/beta hydrolase	−1.09
M1D5S4	Signal recognition particle subunit srp72	−1.10
M1CMY9,M1CMZ2, M1CMZ3,M1CMZ4, M1CMZ5	Superoxide dismutase	−1.25
M1B216,M1B217,M1B218, P49039	Sucrose synthase 2	−1.60

group, eight were enzymes involved in carbohydrate and primary metabolism. Two different alpha-1,4 glucan phosphorylase isoforms (M1ACZ6 and P04045) are involved in the degradation of starch to glucose. Sucrose synthase (M1B216) catalyzes the formation of sucrose, whereas fructose-bisphosphate aldolase (M0ZU27) is an important glycolytic enzyme, and dihydrolipoyl dehydrogenase (M1CL86) is part of the pyruvate dehydrogenase complex. The two enzymes, glucose-6-phosphate 1-dehydrogenase (M1BGQ5) and ribose-5-phosphate isomerase (M1BPR5), are both part of the pentose phosphate pathway. Finally, there was an ammonium transporter (M1BM79).

A comparison of these two groups of proteins indicates that the differentially abundant proteins that were more abundant in the high-yield quotient group are more anabolic in their character, whereas the differentially abundant proteins with lower abundances in the high-yield quotient group are more catabolic. This might indicate a greater ability to utilize the longer light hours at high latitudes among the varieties in the high-yielding group. Another possibility is that the genotypes that yield better in the north are better to grow at lower temperature since low temperature has been shown to induce catabolic proteins (Kosova et al., 2011, 2013).

Except for the proteins involved in primary metabolism, there was no other distinct group among the proteins with low abundances in the high-yield quotient group. Among the differentially abundant proteins with high abundance in the high-yield group, there were several tubulins and seven proteins involved in different types of stress response. TSI-1 protein (M0ZMA8) is a stress-induced protein that can induce PR-proteins, whereas osmotin (M0ZTN0) and P69F protein (M0ZX98) are both PR proteins themselves. The ribonuclease (M1B4H0) is induced in response to abiotic stress. Glutathione *S*-transferase/peroxidase (M1B668) catalyzes the conjugation of GSH to xenobiotics. Allene oxide synthase 2 (M0ZKX2) is involved in the biosynthesis of defensive compounds and jasmonates (Pajerowska-Mukhtar et al., 2008), whereas anthranilate N-benzoyltransferase protein (M1BC65) catalyzes a step in phytoalexin synthesis (Reinhard and Matern, 1989).

The increase in the sugars and sugar alcohols glucose, fructose, fucose, xylose and inositol in the high yielding plants can be interpreted as a sign of higher metabolic activity in these plants. This is consistent with the increased anabolic character of the more abundant proteins in the high-yielding group. Specifically, UDP-glucose 6-dehydrogenase is the enzyme that catalyzes the conversion of UDP-glucose to UPD-glucuronic acid, which is then used in cell wall synthesis (Tenhaken and Thulke), and the increased glucose is the starting material in this process. Xylose is the main component of xylene, which is an important cell wall polysaccharide (Curry et al., 2023). This indicates increased levels of cell wall synthesis in the high-yielding group.

One of the chaperones that are more abundant in the high-yielding genotypes is HSP90, which is down-regulated by cold stress (Kosova et al., 2011). Cold treatment also results in increased abundance of UDP-glucose pyrophosphorylase, which in turn leads to increased production of UDP-glucose (Kosova et al., 2011). We observe higher abundance of UDP-glucose 6-dehydrogenase in the high-yielding group, which can be expected to lead to decreased UDP-glucose. The enzymes regulating UDP-glucose levels have not been observed to change in response to cold stress in potato (Li et al., 2021).

In the metabolomic analysis of the yield groups the high yield group is characterized by lower abundance of flavonoid-glycosides, higher abundance of lyso-phospholipids and higher abundance of sugars. The lower abundance of flavonoid-glycosides is interesting since these compounds also had a higher abundance as a result of cultivation at a higher latitude. A possible interpretation of this is that production of flavonoid-glycosides is a response to extreme long days at the higher latitude, and that the high-yielding plants are better adapted to high latitude, resulting in a relatively lower production of flavonoid-glycosides.

In summary, the biosynthetic nature of a subset of the proteins more abundant in the high-yield group together with the indications of increased membrane remodeling and cell wall synthesis might indicate that the high-yielding genotypes are better able to utilize the increased day length at the northern site. However, the yield differences were relatively small in this experiment and this part of the study would benefit from being repeated more years in order to generate more robust data.

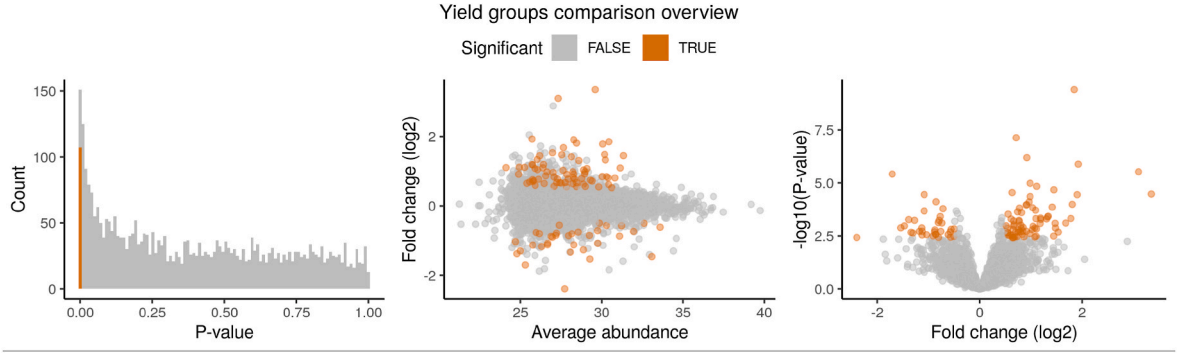

Fig. 6. Illustration of the difference between the yield groups within southern Sweden in 2016 when comparing groups of varieties with comparably higher and lower yields in northern and southern Sweden in 2016. Proteins marked in orange are found in significantly different abundance between these groups (FDR <0.05).

4. Conclusion

This article describes the first extensive field potato proteomics dataset from potatoes grown at high latitudes. Thus, it will provide an important basis for future work and is one of the few crop proteomics datasets from field-grown materials.

We observed small but consistent differences between potato leaves from northern and southern Sweden. This is influenced by that the potato plants in northern Sweden receive, on average, four more hours of light per day during the growing season (20 h). The correlation between protein and mRNA abundance was quite strong in this dataset, which opens up the potential for identifying and using RNA-based biomarkers. However, it is important to remember that patterns of expression should be confirmed in the specific genotypes of interest for breeders before proceeding with future work. This observation was recently also confirmed in another field-based "-omics" study of potato (Boutsika et al., 2023). Proteins with different abundances between our northern and southern field sites were mainly those involved in amino acid metabolism, especially the metabolism of GABA, redox metabolism, hydrolytic activity, and GST activity. When varieties that yield relatively well in northern Sweden were compared to varieties with lower yields, the major difference was that the higher-yielding plants appeared to have more biosynthetic enzymes and stress-related proteins, whereas the lower-yield varieties had more enzymes involved in carbohydrate metabolism.

Our results show that the combination of different types of -omics data can improve our abilities to interpret the differences in the metabolic state of the plants from two locations and to determine which of the many metabolic pathways might be differentially regulated. Our data create a base of information for potato "field-omics," which can improve our understanding of physiological and molecular processes in field-grown plants. Finally, our data indicate that potato is not generally stressed by extremely long days, supported by the limited overlap of proteins identified in our study and those identified in an earlier study of cold included potato proteomics (Li et al., 2021).

CRediT author statement

Svante Resjö;: Conceptualization, Investigation, Visualization, Writing - Original Draft. **Jakob Willforss:** Software, Formal analysis, Visualization, Writing - Original Draft. **Annabel Large:** Formal analysis. **Valentina Siino:** Investigation. **Erik Alexandersson:** Conceptualization, Data curation, Writing - Review & Editing. **Fredrik Levander:** Formal analysis, Data curation, Writing - Review & Editing, Supervision. **Erik Andreasson:** Conceptualization, Writing - Review & Editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The proteomics and transcriptomics data has been deposited in public data repositories. The metabolomics data is available upon request.

Acknowledgements

Support for the proteomic analysis from the Swedish National Infrastructure for Biological Mass Spectrometry (BioMS) and Proteoforms@LU is gratefully acknowledged. Swedish Metabolomics Centre, Umeå, Sweden (www.swedishmetabolomicscentre.se) is acknowledged for performing metabolic profiling using GCMS and LCMS. Sequencing was performed by the SNP & SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI), Sweden, and a Science for Life Laboratory. The SNP & SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. The research was supported by The Swedish Foundation for Strategic Environmental Research (Mistra Biotech), NordPlant [NordForsk (NordPlant #84597)], SLU, Carl Tryggers Stiftelse (CTS16:11), Formas (2020-01211), and PlantLink bioinformatics support. We would like to thank Ganapathi Varma Saripella and Sandeep Kumar Kushwaha for help with the RNA-seq preprocessing, Leli Li for help with the RNA-seq data deposition and Mia Mogren for her excellent technical support. Finally, we would like to thank Zuzanna Sadowska for contributing the potato plant drawing used in fig. 1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphy.2024.109032.

References

Abreha, K.B., Alexandersson, E., Resjo, S., Lankinen, A., Sueldo, D., Kaschani, F., et al., 2021. Leaf apoplast of field-grown potato analyzed by quantitative proteomics and activity-based protein profiling. Int. J. Mol. Sci. 22 (21) https://doi.org/10.3390/ijms222112033.

Alexandersson, E., Jacobson, D., Vivier, M.A., Weckwerth, W., Andreasson, E., 2014. Field-omics-understanding large-scale molecular data from field crops. Frontiers in plant science 5, 286. https://doi.org/10.3389/fpls.2014.00286.

Alexandersson, E., Kushwaha, S., Subedi, A., Weighill, D., Climer, S., Jacobson, D., et al., 2020. Linking crop traits to transcriptome differences in a progeny population of tetraploid potato. BMC Plant Biol. 20 (1). ARTN 12010.1186/s12870-020-2305-x.

Ballman, K.V., Grill, D.E., Oberg, A.L., Therneau, T.M., 2004. Faster cyclic loess: normalizing RNA arrays via linear models. Bioinformatics 20 (16), 2778–2786. https://doi.org/10.1093/bioinformatics/bth327.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30 (15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Boutsika, A., Michailidis, M., Ganopoulou, M., Dalakouras, A., Skodra, C., Xanthopoulou, A., et al., 2023. A wide foodomics approach coupled with metagenomics elucidates the environmental signature of potatoes. iScience 26 (1), 105917. https://doi.org/10.1016/j.isci.2022.105917.

Carpentier, S.C., Witters, E., Laukens, K., Deckers, P., Swennen, R., Panis, B., 2005. Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. Proteomics 5 (10), 2497–2507. https://doi.org/10.1002/pmic.200401222.

Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., et al., 2012. A cross-platform toolkit for mass spectrometry and proteomics. Nat. Biotechnol. 30 (10), 918–920. https://doi.org/10.1038/nbt.2377.

Curry, T.M., Pena, M.J., Urbanowicz, B.R., 2023. An update on xylan structure, biosynthesis, and potential commercial applications. Cell Surf 9, 100101. https://doi.org/10.1016/j.tcsw.2023.100101.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29 (1), 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Eriksson, D., Carlson-Nilsson, U., Ortiz, R., Andreasson, E., 2016. Overview and breeding strategies of table potato production in Sweden and the fennoscandian region. Potato Res. 59 (3), 279–294. https://doi.org/10.1007/s11540-016-9328-6.

Ewels, P., Magnusson, M., Lundin, S., Kaller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32 (19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

Food and Agriculture Organization of the United Nations. Faostat - data [Online]. Available: http://www.fao.org/faostat/en/#data. (Accessed 15 September 2020).

Gullberg, J., Jonsson, P., Nordstrom, A., Sjostrom, M., Moritz, T., 2004. Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of Arabidopsis thaliana samples in metabolomic studies with gas chromatography/mass spectrometry. Anal. Biochem. 331 (2), 283–295. https://doi.org/10.1016/j.ab.2004.04.037.

Gullner, G., Komives, T., Kiraly, L., Schroder, P., 2018. Glutathione S-transferase enzymes in plant-pathogen interactions. Front. Plant Sci. 9, 1836. https://doi.org/10.3389/fpls.2018.01836.

Hakkinen, J., Vincic, G., Mansson, O., Warell, K., Levander, F., 2009. The proteios software environment: an extensible multiuser platform for management and analysis of proteomics data. J. Proteome Res. 8 (6), 3037–3043. https://doi.org/10.1021/pr900189c.

Hardigan, M.A., Crisovan, E., Hamilton, J.P., Kim, J., Laimbeer, P., Leisner, C.P., et al., 2016. Genome reduction uncovers a large dispensable genome and adaptive role for copy number variation in asexually propagated Solanum tuberosum. Plant Cell 28 (2), 388–405. https://doi.org/10.1105/tpc.15.00538.

Haverkort, A.J., 1990. Ecology of potato cropping systems in relation to latitude and altitude. Agric. Syst. 32 (3), 251–272. https://doi.org/10.1016/0308-521X(90)90004-A.

Jaakola, L., Hohtola, A., 2010. Effect of latitude on flavonoid biosynthesis in plants. Plant Cell Environ. 33 (8), 1239–1247. https://doi.org/10.1111/j.1365-3040.2010.02154.x.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., et al., 2014. Erratum to: euro-cordex: new high-resolution climate change projections for European impact research. Reg. Environ. Change 14 (2), 579–581. https://doi.org/10.1007/s10113-014-0587-y.

Kato, Y., Murakami, S., Yamamoto, Y., Chatani, H., Kondo, Y., Nakano, T., et al., 2004. The DNA-binding protease, CND41, and the degradation of ribulose-1,5-bisphosphate carboxylase/oxygenase in senescent leaves of tobacco. Planta 220 (1), 97–104. https://doi.org/10.1007/s00425-004-1328-0.

Kim, S., Pevzner, P.A., 2014. MS-GF+ makes progress towards a universal database search tool for proteomics. Nat. Commun. 5, 5277. https://doi.org/10.1038/ncomms6277.

Kopylova, E., Noe, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28 (24), 3211–3217. https://doi.org/10.1093/bioinformatics/bts611.

Kosova, K., Vitamvas, P., Prasil, I.T., Renaut, J., 2011. Plant proteome changes under abiotic stress–contribution of proteomics studies to understanding plant stress response. J Proteomics 74 (8), 1301–1322. https://doi.org/10.1016/j.jprot.2011.02.006.

Kosova, K., Vitamvas, P., Planchon, S., Renaut, J., Vankova, R., Prasil, I.T., 2013. Proteome analysis of cold response in spring and winter wheat (Triticum aestivum) crowns reveals similarities in stress adaptation and differences in regulatory processes between the growth habits. J. Proteome Res. 12 (11), 4830–4845. https://doi.org/10.1021/pr400600g.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9 (4), 357–359. https://doi.org/10.1038/nmeth.1923.

Lankinen, A., Abreha, K.B., Masini, L., Ali, A., Resjo, S., Andreasson, E., 2018. Plant immunity in natural populations and agricultural fields: low presence of pathogenesis-related proteins in Solanum leaves. PLoS One 13 (11), e0207253. https://doi.org/10.1371/journal.pone.0207253.

Li, H., Luo, W., Ji, R., Xu, Y., Xu, G., Qiu, S., et al., 2021. A comparative proteomic study of cold responses in potato leaves. Heliyon 7 (2), e06002. https://doi.org/10.1016/j.heliyon.2021.e06002.

Liljeroth, E., Bengtsson, T., Wiik, L., Andreasson, E., 2010. Induced resistance in potato to Phytphthora infestans—effects of BABA in greenhouse and field tests with different potato varieties. Eur. J. Plant Pathol. 127 (2), 171–183. https://doi.org/10.1007/s10658-010-9582-4.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15 (12), 550. https://doi.org/10.1186/s13059-014-0550-8.

McLoughlin, F., Basha, E., Fowler, M.E., Kim, M., Bordowitz, J., Katiyar-Agarwal, S., et al., 2016. Class I and II small heat shock proteins together with HSP101 protect protein translation factors during heat stress. Plant Physiol 172 (2), 1221–1236. https://doi.org/10.1104/pp.16.00536.

Michaeli, S., Fromm, H., 2015. Closing the loop on the GABA shunt in plants: are GABA metabolism and signaling entwined? Front. Plant Sci. 6, 419. https://doi.org/10.3389/fpls.2015.00419.

Molmann, J.A.B., Dalmannsdottir, S., Hykkerud, A.L., Hytonen, T., Samkumar, A., Jaakola, L., 2021. Influence of Arctic light conditions on crop production and quality. Physiol Plant 172 (4), 1931–1940. https://doi.org/10.1111/ppl.13418.

Pajerowska-Mukhtar, K.M., Mukhtar, M.S., Guex, N., Halim, V.A., Rosahl, S., Somssich, I. E., et al., 2008. Natural variation of potato allene oxide synthase 2 causes differential levels of jasmonates and pathogen resistance in Arabidopsis. Planta 228 (2), 293–306. https://doi.org/10.1007/s00425-008-0737-x.

Perez-Riverol, Y., Bai, J., Bandla, C., Garcia-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., et al., 2022. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 50 (D1), D543–D552. https://doi.org/10.1093/nar/gkab1038.

Polpitiya, A.D., Qian, W.J., Jaitly, N., Petyuk, V.A., Adkins, J.N., Camp 2nd, D.G., et al., 2008. DAnTE: a statistical tool for quantitative analysis of -omics data. Bioinformatics 24 (13), 1556–1558. https://doi.org/10.1093/bioinformatics/btn217.

Ramos-Ruiz, R., Martinez, F., Knauf-Beiter, G., 2019. The effects of GABA in plants. Cogent Food Agric. 5 (1), 167055310.1080/23311932.2019.1670553.

Rappsilber, J., Mann, M., Ishihama, Y., 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. 2 (8), 1896–1906. https://doi.org/10.1038/nprot.2007.261.

Reinhard, K., Matern, U., 1989. The biosynthesis of phytoalexins in Dianthus caryophyllus L. cell cultures: induction of benzoyl-CoA:anthranilate N-benzoyltransferase activity. Arch. Biochem. Biophys. 275 (1), 295–301. https://doi.org/10.1016/0003-9861(89)90376-7.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., et al., 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43 (7), e47 https://doi.org/10.1093/nar/gkv007.

Roitsch, T., Himanen, K., Chawade, A., Jaakola, L., Nehe, A., Alexandersson, E., 2022. Functional phenomics for improved climate resilience in Nordic agriculture. J. Exp. Bot. 73 (15), 5111–5127. https://doi.org/10.1093/jxb/erac246.

Rojas, C.M., Senthil-Kumar, M., Tzin, V., Mysore, K.S., 2014. Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. Front. Plant Sci. 5, 17. https://doi.org/10.3389/fpls.2014.00017.

Schauer, N., Steinhauser, D., Strelkov, S., Schomburg, D., Allison, G., Moritz, T., et al., 2005. GC-MS libraries for the rapid identification of metabolites in complex biological samples. FEBS Lett. 579 (6), 1332–1337. https://doi.org/10.1016/j.febslet.2005.01.029.

Schumacher, C., Krannich, C.T., Maletzki, L., Kohl, K., Kopka, J., Sprenger, H., et al., 2021. Unravelling differences in candidate genes for drought tolerance in potato (Solanum tuberosum L.) by use of new functional microsatellite markers. Genes 12 (4). https://doi.org/10.3390/genes12040494.

Shi, X., Tian, Z., Liu, J., van der Vossen, E.A., Xie, C., 2012. A potato pathogenesis-related protein gene, StPRp27, contributes to race-nonspecific resistance against Phytophthora infestans. Mol. Biol. Rep. 39 (2), 1909–1916. https://doi.org/10.1007/s11033-011-0937-5.

Sprenger, H., Erban, A., Seddig, S., Rudack, K., Thalhammer, A., Le, M.Q., et al., 2018. Metabolite and transcript markers for the prediction of potato drought tolerance. Plant Biotechnol. J. 16 (4), 939–950. https://doi.org/10.1111/pbi.12840.

Teleman, J., Dowsey, A.W., Gonzalez-Galarza, F.F., Perkins, S., Pratt, B., Rost, H.L., et al., 2014. Numerical compression schemes for proteomics mass spectrometry data. Mol. Cell. Proteomics 13 (6), 1537–1542. https://doi.org/10.1074/mcp.O114.037879.

Teleman, J., Chawade, A., Sandin, M., Levander, F., Malmstrom, J., 2016. Dinosaur: a refined open-source peptide ms feature detector. J. Proteome Res. 15 (7), 2143–2151. https://doi.org/10.1021/acs.jproteome.6b00016.

Tenhaken, R., and Thulke, O. Cloning of an Enzyme that Synthesizes a Key Nucleotide-Sugar Precursor of Hemicellulose Biosynthesis from Soybean: UDP-Glucose Dehydrogenase. (32-889 (Print)).

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., et al., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7 (3), 562–578. https://doi.org/10.1038/nprot.2012.016.

Willforss, J., Chawade, A., Levander, F., 2019. NormalyzerDE: online tool for improved normalization of omics expression data and high-sensitivity differential expression analysis. J. Proteome Res. 18 (2), 732–740. https://doi.org/10.1021/acs.jproteome.8b00523.

Willforss, J., Siino, V., Levander, F., 2021. OmicLoupe: facilitating biological discovery by interactive exploration of multiple omic datasets and statistical comparisons. BMC Bioinf. 22 (1), 107. https://doi.org/10.1186/s12859-021-04043-5.

Yoshida, Y., Umeno, A., Shichiri, M., 2013. Lipid peroxidation biomarkers for evaluating oxidative stress and assessing antioxidant capacity in vivo. J. Clin. Biochem. Nutr. 52 (1), 9–16. https://doi.org/10.3164/jcbn.12-112.