Introduction and goals

After rice and wheat, potato is the world's third most important staple food. Potato is a cosmopolite plant which can be grown in tropical areas up to the Arctic region. With 1,500-2,000 species, the *Solanum* is the largest genus in the *Solanaceae* family. However, wild potato species are grown mainly in their native Middle- and South-American regions, in most parts of the world, *S. tuberosum* is cultivated as potato. The cultivated potato has a large secondary gene pool of related wild species, from which, by crossing or somatic hybridisation genes can be introgressed into *S. tuberosum*.

At the Potato Research Centre (PRC) at Keszthely a potato resistance breeding program was started about 60 years ago, and now there are more than a dozen varieties which harbor different resistance genes. However, the early pedigree of the breeding lines is not available, it is known that lines with introgressed resistance genes have been obtained in the frame of international cooperation, and that same germplasm was used for breeding in many other countries, too. In the breeding material used at the PRC, genes conferring resistance against PVY, PVX, PVS, PVA and PVV viruses, against late blight, nematodes, and potato scab are present, among others. Potato cultivation generally requires the intensive use of agrochemicals. Hence the use of resistant cultivars would contribute to significant cost reduction for the farmers, alleviate the chemical pressure on the environment and to healthier food for the consumers.

The main goal of the present project was to explore the genetic bases of stress resistance in potatoes by applying high-throughput molecular genetic technologies. To this end, three different approaches have been used: whole genome sequencing, transcriptome sequencing, and microarray genotyping.

The purpose of the whole genome sequencing was to construct a haploid-chromosome resolved whole genome sequence (WGS) of the tetraploid potato that can be used then as a reference tool for the localisation of genes of interest present in our breeding material and also for the isolation of genes present in the genome-sequenced line. Here, we sequenced and reconstructed the WGS of the Keszthely bred potato variety White Lady, which contains several isolated and not yet isolated resistance genes.

The transcriptome sequencings aimed to analyse pathogen infection triggered gene expressional changes and screen the transcripts for motifs typical for resistance genes. By these approaches, we believe we can identify resistance gene candidates. We focused on the two main pathogens of potatoes, namely, the PVY virus and the late blight disease causing *Phytophthora infestans*. For PVY plant materials conveying four extreme resistance genes, the Ry_{sto}, Ry_{chc}, Ry_{hou} and Ry_{adg} originating from the wild tuber-bearing potato species, the *Solanum stoloniferum, S. chacoense, S. hougasii* and *S. tuberosum* subsp. *andigenum*, respectively, have been used. For *P. infestans* resistance analysis, plant materials containing strain-specific resistance genes from *S. demissum* and resistance genes from *S. bulbocastanum* and *S. hougasii* have been used.

A further purpose of the transcriptome sequencings was to identify water shortage-induced expressional changes in different potato genotypes to understand the molecular genetic background of drought tolerance in potatoes.

The microarray technology was applied to localise the analysed resistance genes with SNP (single nucleotide polymorphism) markers. The used Potato V4 beadchip contains 31,190 SNP markers, giving a highly saturated map of the potato genome. Our purpose was by the use of segregating populations to identify closely linked markers to a given trait.

Results of whole genome sequencing of the variety White Lady

In the first step, whole genome sequencing (WGS) of the cultivar White Lady (WL) was performed on a short read (150 bp) producing platform (NextSeq 500, Illumina) at our laboratory. For assembling at first a *S. tuberosum* subs. *andigenum* sequence (ADG1) (Kyriakidou et al., 2020) was used. Although the reconstructed WL genome was useful as a reference genome for transcriptome alignments, it had many gaps. Recently, a haplotype-resolved tetraploid potato (*S. tuberosum*) genome sequence of the genotype C88 was published (Bao et al., 2022), which we used as a reference and for the reconstruction of a chromosome-scale haplotype-resolved whole genome sequence of the variety White Lady (Table 1. and 2.). Our colleague, Krisztián Frank (MATE), has filtered and assembled the short read experiments. Although, even with this approach, gaps and repeated sequences remained and caused complications.

Table 1. Statistical data of the WL genome sequence reconstructed from short reads

	Phase 1	**Phase 2**	**Phase 3**	**Phase 4**
Number of chromosomes	12	12	12	12
Unmapped scaffolds in reference	352	110	215	178
Mapped scaffolds	3 947 834	4 253 016	3 939 486	4 138 584
Length of mapped scaffolds (bp)	1 006 022 039	1 066 439 709	1 006 311 196	1 050 196 092
Unlocalised fragments	1 562 486	1 556 306	1 545 026	1 559 575
Length of unlocalised fragments (bp)	405 577 940	405 768 443	401 819 366	404 853 695
Unplaced fragments	2 553 848	2 307 018	2 556 249	2 381 583
Length of unplaced fragments (bp)	557 585 414	505 450 346	557 507 972	516 276 236
Length of gaps (bp)	394 783 400	425 301 600	393 948 600	413 858 400

Table 2. Haplotype-resolved genome size of the variety White Lady, reconstructed from short reads

Chromosomes	**Phase 1 (bp)**	**Phase 2 (bp)**	**Phase 3 (bp)**	**Phase 4 (bp)**
Chromosome 1	136 480 167	172 103 441	168 233 317	175 344 038
Chromosome 2	89 197 929	96 065 508	96 452 861	95 861 675
Chromosome 3	119 911 202	114 144 583	120 492 600	114 768 193
Chromosome 4	107 885 409	140 162 514	87 214 681	107 402 602
Chromosome 5	116 269 889	108 576 713	113 543 973	113 244 825
Chromosome 6	126 382 539	134 901 792	135 198 105	139 096 851
Chromosome 7	102 676 328	91 478 098	99 113 596	101 932 565
Chromosome 8	111 901 174	142 504 527	124 538 152	125 755 955
Chromosome 9	131 014 616	131 711 152	88 447 919	129 286 689
Chromosome 10	113 968 232	114 976 769	118 982 204	117 735 158
Chromosome 11	94 357 420	103 926 496	100 636 504	98 517 456
Chromosome 12	136 468 143	135 369 484	129 896 781	137 821 743
	1 386 513 048	1 485 921 077	1 382 750 693	1 456 767 750

Note: Haplotypes are indicated here as phase 1-4.

The haplotype-resolved genome sizes obtained after assembling the short reads are significantly larger than the potato's 844 Mbp suggested genome size. However, Kyriakidou et al. (2020), reported that among the six assembled potato genomes, the genome size of *S. curtilobum* was above 1.2 Gbp, to clarify possible contig redundancy and to clear the problem of repeated sequence regions as well as to fill the gaps in the genome, we decided to re-sequence the WL genome with long reads on a PacBio Sequel Ile platform by using a SMRTcell. This experiment was implemented at the Szentágothai Research Centre (SzRC), Pécs, with the cooperation of Péter Urbán. Bence Gálik (SzRC) had filtered, and preliminary arranged the reads on their SMRTLink v.11.1 server. Row data of the long reads sequencing are shown in Table 3.

Table 3. Long read row data of the whole genome sequencing of the variety White Lady

Value	Analysis Metric
365,322,525,216	Polymerase Read Bases
3,204,696	Polymerase Reads
113,996	Polymerase Read Length (mean)
247,750	Polymerase Read N50
13,840	Longest Subread Length (mean)
26,250	Longest Subread N50
33,709,322,240	Unique Molecular Yield

At ten times genome coverage, the resulted HiFi reads' mean length was above 8000 bp. Only filtered PacBio sequences have been used for contig construction (Table 4). The distribution of the different length contigs is shown in Figure 1.

Table 4. Main results of contig construction.

Contig Type	Polished Contigs	Maximum Contig Length	Mean Contig Length	Median Contig Length	N50 Contig Length	Sum of Contig Lengths	E-size (sum of squares / sum)	Number of Circular Contigs
Primary Contigs	13,558	2,520,239	75,278	33,893	164,326	1,020,620,947	305,916	252

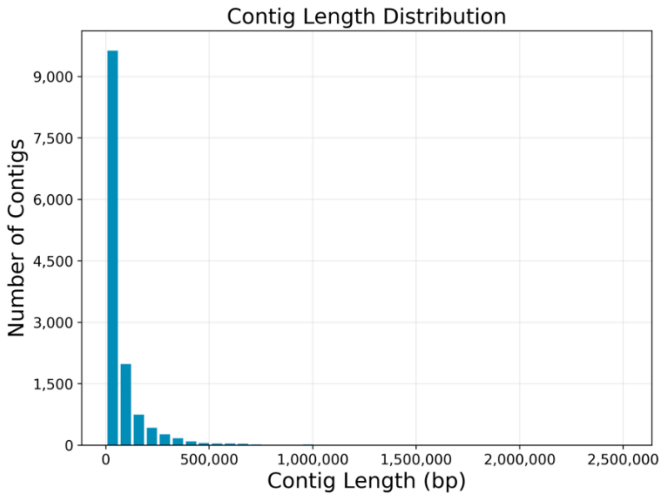

Figure 1. Number of the different length contigs of the White Lady WGS

With the generated high-quality consensus HiFi sequences and reconstructed contigs were used for scaffolding with GAPLESS and LRSCAF programs. By this, the gaps between contigs have been filled. In that process, long read algorithms, like minimap, blasr were used to align the sequences on the genome. As a result of this type of mapping the gaps between contigs, and the 'N'-marked misses in the sequences could be filled. However, by this approach, a chromosome-scale haplotype-resolved

reference guided *de novo* genome sequence of the variety White Lady was created, although it should be annotated before publication. That is the task on which we are working now. It is also important to mention that this WGS contains those genomic DNA sequences, including genes, which are not present on the C88 reference genome.

The transcriptome sequence of the variety White Lady

Besides that, to explore allele-specific gene expression, an IsoSeq transcriptome sequencing with long HiFi reads on the PacBio Sequel Ile platform with a SMRTcell was also performed, using polyA enriched RNA sample. The RNA sample was the same for both the short and long-read sequencing and was extracted from young leaves of a WL plant. Filtering of data and preliminary analysis were done on the SMRTLink v11.1 server. Row data of transcriptome sequencing is shown in Table 5.

Table 5. Row data of long-read transcriptome sequencing of White Lady

Value	Analysis Metric
302,980,041,716	Polymerase Read Bases
3,780,158	Polymerase Reads
80,150	Polymerase Read Length (mean)
183,750	Polymerase Read N50
7,384	Longest Subread Length (mean)
27,250	Longest Subread N50
18,515,859,456	Unique Molecular Yield

With the IsoSeq method, no assembling is required because of their length individual transcripts reliably represent the isoforms/alleles. This approach allows the easy identification of alternative splicing variants and other expressional events, like fused transcripts, non-coding RNAs, etc. With this approach, we think most open reading frames, including allele-specific isoforms, can be identified in the White Lady genome.

Using the IsoSeq pipeline, the obtained transcripts were further studied, and 1,845,714 high-quality sequences (Non-Concatamer Reads with 5' and 3' Primers and Poly-A Tail) were identified (Table 6.).

Table 6. Row data the IsoSeq analysis of the White Lady transcriptome

Value	Analysis Metric
2,200,862	Reads
1,865,797	Reads with 5' and 3' Primers

Value	Analysis Metric
1,851,886	Non-Concatamer Reads with 5' and 3' Primers
1,845,714	Non-Concatamer Reads with 5' and 3' Primers and Poly-A Tail
2,783	Mean Length of Full-Length Non-Concatamer Reads
1	Unique Primers
1,865,797	Mean Reads per Primer
1,865,797	Max. Reads per Primer
1,865,797	Min. Reads per Primer
335,065	Reads without Primers
84.79%	Percent Bases in Reads with Primers
84.77%	Percent Reads with Primers

The IsoSeq analysis 130.274 high-quality transcripts could be reconstructed, which also represent the allelic-isoforms expressed in the tetraploid genome. This number is almost identical to the 130,820 transcripts identified from the short-read transcriptome assembly using the same RNA sample. The distribution of the consensus isoforms is shown in Figure 2.

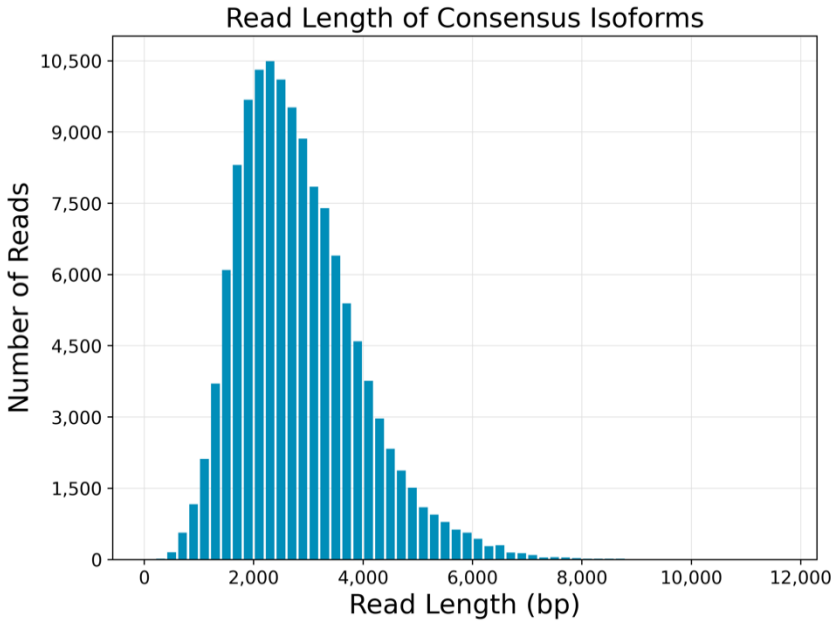

Figure 2. Distribution of isoforms obtained with the Isoseq analysis

In the publication of the chromosome-scale haploid resolved whole genome sequence of the tetraploid potato variety White Lady, the results of the transcriptome sequencing, the identification of nucleotide binding site (NBS) containing genes and the identification of genes with high similarity to cloned

resistance genes will be included. To this end, we are working on the final reconstruction of the WL WGS from the short and long reads, to exclude possible contig redundances, fill the still existing gaps between contigs, and bridge repeated regions by bioinformatic tools. Then we will perform an *ab initio* gene prediction by constructing gene models, paying particular attention to alternative isoforms and new transcript variations. After that, by functional characterisation annotation will be done using the NCBI blast functions and the UNIPROT database. Domain prediction will be made with the InterProScan program, and the already explored NBS containing genes of WL will be identified using the SolariX website (www.cibiv.at/SolariX/). Finally, the results will be aligned with the KEGG reference pathways.

Among the resistance genes in White Lady the Ry_{sto}, Rx2, R1, R2, R3a, and partially the R8 genes have already been isolated. Besides these, by phenotyping and using the Black differential lines, previously, we revealed the presence of the R4, R6, R7, R9, R10 and R11 race-specific *P. infestans* resistance genes in WL. In the present analysis, we use the complete coding sequences of 607 resistance genes presently found in the NCBI for the Solanum genus. Further, we will identify the 562 NBS-containing genes based on the NBS motifs explored in WL by Prakash et al. (2020). It is suggested that these approaches can identify gene candidates and their real function can be proven in further studies.

Transcriptome analysis of P. infestans resistance, of PVY resistance, as well as of drought tolerance

The purpose was to identify late blight resistance genes which are not yet cloned. Besides the R genes of White Lady, it was assumed by the potato breeders that the breeding line J101K3 contains a *P. infestans* resistance gene that originates from *S. bulbocastanum*. Further, they thought also that the varieties Kastia and Sárvári borostyán might contain a *S. hougasii*-derived *P. infestans* resistance gene. It was also assumed that these two varieties have a PVY resistance gene (Ryhou) originating from *S. hougasii*. The breeding line 89.451 contains a PVY extreme resistance gene (Ry_{adg}) originating from *S. tuberosum* subs. *andigena*.

True potato seeds were already available, and the following F1 segregating populations could be grown:

J101K3 (Rpi_{blb}) (S to PVY) x 89.451 (Ry_{adg}) (S to *P. infestans*) – 39 genotypes

Sárvári borostyán (Rpi_{hou}) and (Ry_{hou}) x S440 (S and S) – 72 genotypes

Kastia (Rpi_{hou}) and (Ry_{hou}) x S440 (S and S) – 65 genotypes

89.450 (Ry_{adg}) x Somogyi kifli (S) – 65 genotypes

The assumed resistance is revealed in parenthesis.

Methods of the P. infestans infection tests

For the inoculation with *P. infestans* the isolate MP-1548 was used. This isolate was obtained upon request from Dr. Jadwiga Sliwka, IHAR, Mlochow, Poland. It was collected in 2012 from the potato variety Irys in the Mazowieckie Region. However, regarding the obtained information, it belongs to the A1 mating type, and the type of virulence of it is 1,3,4,6,7,8,10, and 11, we found that this isolate produced oospores, which is not characteristic of the A1 mating type, and its virulence was found to be 1,2,3,4,5,6,7,8,10, and 11 types. All of these indicated that the obtained isolate is a mixed A1-A2 type. Nevertheless, we used this isolate for infection tests because of the shortage of time.

At first, on the parental lines and the segregating populations, detached leaves assay was applied in six repeats to evaluate the level of tolerance/resistance. Lesion diameters were measured on the 4th, 5th, 6th and 7th days after infection. After determining the phenotypes, the parental lines were grown in a phytotron under controlled conditions. Infection with the *P. infestans* isolate was done directly on the abaxial leaf surface of living plants. For each infected leaf, a structure was built at first to keep the leaf in a horizontal position up with its abaxial side. For the infection solution, *P. infestans* was grown on 1.0-1.5 cm thick tuber slices of the variety Hópehely, which have been previously sterilised in 10% sodium hypochlorite solution for 10 minutes and rinsed several times before inoculation. Tuber slices were incubated at 17^{o}C for six days to let the mycelium grow. Then, from 20 tuber slices, the phytophthora mycelium was moved with sterile tools into 10 ml of sterile distilled water and steered vigorously to let the sporangia be separated from the mycelium. After filtering the solution through four sheets of gauze, sporangium density was measured with a Bürker chamber, adjusted to 15,000 sporangium/mL, and incubated for 2 hrs at 5-6^{o}C. Before infection, the solution was left to warm to 20^{o}C. The inoculation was done by applying 50 µL in one drop from the solution on the leaf surface. The infection solution was removed the next morning by pipetting it off. Five plants of each genotype were infected, and two leaves of each plant have used for infection.

Samples for transcriptome analyses were taken at four different time points (18, 24, 48 and 72 hours) after infection from uninfected leaflets of infected leaves.

The samples were collected on dry ice and were stored at -80^{o}C until RNA extraction.

It should be mentioned that because of the discrepancy between the description of the MP-1548 isolate and our observations, with the contribution of József Bakonyi (Plant Protection Institute, Centre for Agricultural Research, HAS), a pure culture of the MP-1548 isolate that is not producing oospores and also a monosporic culture of the isolate has been produced in these days. This monosporic culture does not produce oospores and belongs to A1 mating type. Whole genome sequencing of these three isolate types will be performed soon to explore their differences.

Methods of the PVY infection tests

For the PVY resistance tests, the PVY-NTN-H isolate was used. The virus was maintained on a susceptible host, tobacco. Infected tobacco leaves were harvested, chopped and mashed in a mortar with phosphate buffer. The liquid was filtered through two sheets of gauze. Half of the adaxial surface of the largest leaflet of a leaf on a young potato plant was sprinkled with carborundum powder, and the virus solution was rubbed into the leaf using a glass spatula. Leaves of young plants with about six leaves of the parental lines and the segregating populations were infected, and the symptoms were observed. After six weeks, uninfected leaflets were sampled, and a DAS-ELISA test was performed to determine the resistance of the genotypes.

To detect the virus by DAS-ELISA, the AP-PAbs (alkalic-phosphatase conjugated PAbs9 antibody) (Loewe Biochemica GmbH, Germany) was used with p-Nitrofenil-phosphate and polystyrol plate (Corning 96 Well EIA/RIA Clear Flat Bottom Polystyrene High Band Microplate, USA). The experiment was performed by following the protocol of the producer. *Nicotiana tabacum* Xanthi-nc leaf juice was used for positive control, and sterile, in vitro potato plants were used for negative control. The extinction values were measured at 405 nm on a BMG Labtech Spectrostar Nano photometer (BMG Labtech, Germany).

From the parental lines, samples for transcriptome analyses were taken at six different time points (2, 4, 6, 24, 48 and 72 hours) after infection and one sample for control immediately before infection. The samples taken were from uninfected leaflets of infected leaves.

The leaf samples were collected on dry ice and were stored at -80°C until RNA extraction.

Methods of drought tolerance tests

Two mid-late varieties, Hópehely and Démon have been used for exploring the water shortage-induced gene expressional changes. Potato breeders have previously identified these two varieties to give different reactions to dry periods on the field.

The experiments were established in a greenhouse. Fifty kg of soil : peat (1 : 1) mixture was filled into pots, and 16 replications were used in a randomised complete block design. Under a controlled environment (day/night temperature 25 / 21 °C, 50% relative humidity, 18 h photoperiod), two water levels, 80% water holding capacity for control, and 50% water holding capacity for drought stress were used. Water was imposed 18 days after tuber planting. This experiment was implemented by a PhD student, Waqar Nasir, and his supervisor, Zoltán Tóth, at the Georgikon Campus.

For transcriptome analysis, leaves were sampled weekly, and on the 36th, 54th, 72nd, and 90th days, when four-four replications per treatment were harvested and measured besides leaves, root samples were also taken. The samples were collected on dry ice and were stored at -80°C until RNA extraction.

Wetlab methods of transcriptome analyses

For transcriptomic analyses, RNA was extracted from the leaf samples with Trizol and sequencing libraries were prepared with the TruSeq RNA Library Prep kit v2 (Illumina, USA) according to the producer's guidelines. Sequencings were done on a NextSeq 500 type sequencer.

In the case of late blight resistance, the control and the four treated samples of White Lady, Sárvári borostyán and Kastia were transcriptome sequenced. These are, in total, 15 transcriptomes.

In the case of PVY resistance, the control and the 24th, 48th and 72nd hrs samples after treatment of White Lady, Somogyi kifli, S440, Kastia, Pannónia and the breeding line 89.451 were transcriptome sequenced. These are, in total, 24 transcriptomes.

For water shortage response, the control and the treated plants were sampled on the 36th, 54th, 72nd, and 90th days after the start of treatment. In this experiment, two varieties, the Hópehely and Démon, which show different reactions to drought, were used and transcriptome sequenced. These are, in total, 16 transcriptomes.

Transcriptome construction and analysis methods

De novo reconstruction of the transcriptomes was done using the SOAPdenovo program.

For gene expression analyses, the *de novo* transcriptome of White Lady was used as a reference. The "index" and "quant" commands of the Salmon program package were used to quantify RNA-Seq data. Quantified transcripts were analysed and visualised with the DESeq2 package in 'R' environment.

Further analyses of the transcript sequences are done with the seqkit program, supplemented with the "grep", "awk", "join", and "comm" commands.

Results of transcriptome analyses

Results of the late blight resistance analyses

The segregations for late blight resistance were as follows:

J101K3 (*Rpi**blb***) x 89.451 (S to *P. infestans*) – R : S = 1 : 38

Sárvári borostyán (*Rpi**hou***) x S440 (S) – R : S = 6 : 66

Kastia (*Rpi**hou***) x S440 (S) – R : S = 8 : 57

Interestingly, lesions occurred on the J101K3 line after 48 hrs, and after seven weeks, more than half of the infected leaves turned brown. These results and the segregation ratio indicate that J101K3 has a different type of *P. infestans* resistance than the hypothesised broad-spectrum resistance, conveyed by some already cloned Rpi genes originating from *S. bulbocastanum*. Symptoms of 89.451 after seven weeks were complete leaf browning for all infected plants and stem browning at two of the five infected plants. In this cross combination, only a single F1 genotype was identified, which had no symptoms in the detached leaf assay. The reasons for that will be analysed further.

P. infestans infected Sárvári borostyán leaves showed lesions after 48 hrs, but the spread of the infection stopped, and after six weeks, the plants grew healthy. It was concluded that Sárvári borostyán has some late blight resistance but not a broad-spectrum type. S440 had moderate symptoms, which occurred on the 6th day after infection. Kastia, which was assumed to contain a *Rpi* gene, showed late blight symptoms, and after five weeks besides the leaves, the stems browned for four out of the five infected plants. It is concluded that Kastia is somewhat sensitive to late blight, although the presence of race-specific resistance cannot be excluded.

Reasons for the observed six and eight symptomless F1 genotypes in the later two cross-combinations need further investigation. One reason can be that the detached leaves assay was terminated after the 7th day, and symptoms would appear just after that, but among others, the reasons can also be genetic. Next, we repeat the infection tests on living plants with these genotypes and some sensitive siblings.

To explore the R genes in the variety White Lady five plants have been infected with the MP-1548 isolate. Lesions were observed at the infection spot, but the plants grew healthy after seven weeks of inoculation.

Until now, the White Lady's transcriptomes (one control and four treated (18, 24, 48 and 72 hrs after infection) were reconstructed. In control, 92,079 transcripts were obtained. In total, 2,198, 845, 1,100 and 1,210 transcripts showed significant upregulation 18, 24, 48 and 72 hours after infection, respectively. Since the transcriptomes are not yet completely annotated, we have chosen the most upregulated first 200 transcripts and checked them one by one in the NCBI and then in the UniProt database to identify the resistance gene-like sequences among them.

In these five transcriptomes, sequence similarity search with the Solanum resistance gene collection (607 genes) is just running on our server, and also with the 562 NBS motifs (SolariX database), which are characteristic for resistance genes and have been identified in the variety White Lady we will soon start an analysis.

Results of the PVY resistance analyses

PVY infections were done on three plants for each genotype. The resistant to sensitive segregation ratios for the investigated population are as follows:

Sárvári borostyán (*Ry**hou***) x S440 (S) – R : S = 30 : 42

Kastia (*Ry**hou***) x S440 (S) – R : S = 49 : 16

89.450 (*Ry**adg***) x Somogyi kifli (S) – R : S = 36 : 29

The F1 genotypes for the Sárvári borostyán, and Kastia here were the same as for late blight resistance analysis.

Segregation ratios of these three populations indicate the presence of a PVY resistance gene in the resistant parent. R means that neither of the three replications of each genotype showed any symptoms of infection, and also, by DAS-ELISA the virus could not be detected in the R genotypes. S genotypes showed typical symptoms of PVY infection. We aim to identify new PVY extreme resistance gene candidates by transcriptomic approach. Because this and the population sizes are not too large, we do not enter into statistical evaluations at this stage.

RNA-sequencings were performed on the White Lady, Somogyi kifli, S440, Kastia, Pannónia and 89.451 genotypes using the 24th, 48th and 72nd hrs samples and the control of each. From these 20 transcriptomes, for the moment, eight, the 89.451 and Somogyi kifli were assembled, and evaluation was started. The Ryadg gene conveys PVY extreme resistance in 89.451. However, closely linked markers are published for the *Ry**adg***, but the gene is not yet cloned. Besides the haplotype-resolved WGS, in this project, our primary aim is to identify gene candidates for *Ry**adg***, and then to analyse the presence/absence of the gene candidates in the segregating population. After publishing those results, the real function of the gen candidate(s) will be analysed. We are currently screening the transcriptomes with an expressional and a sequence similarity + NBS motif approaches, as described at the late blight resistance. For PVY extreme resistance, the constitutively expressed *Ry**sto*** gene (White Lady), and the *Ry**chc*** (Pannónia) have been recently isolated. The White Lady and Pannónia transcriptomes will be used to explore what kind of PVY infection-induced expressional changes can be observed in the presence of these genes, and whether this information can be utilised in the identification and/or expressional profiling of *Ry**adg***.

Results of drought tolerance tests

As indicated in the methods part, transcriptome sequencing of the leaf samples taken at the four harvesting periods was performed. We planned to sequence also the root samples, but the quality of the extracted RNA was insufficient for sequencing.

Nasir and Tóth (2021) have published the analysed morphological and physiological characteristics. The generated 16 transcriptomes are assembled, and their analysis is processed.

Microarray analysis

Methods of microarray analysis

The Potato V4.0 beadchip (Illumina, USA) was used for microarray analysis. This improved chip contains 31.190 SNP-s.

For marker identification, the plant material was mainly as in the previous experiments, but for populations with a slightly higher number of genotypes:

J101K3 x 89.451 – 40 genotypes

Sárvári borostyán x S440 – 80 genotypes

Kastia x S440 – 65 genotypes

89.450 x Somogyi kifli – 65 genotypes

For genotyping, 38 genotypes containing different varieties, breeding lines and wild potato accessions for the gene collection of the PRC have been used.

DNA was extracted with the NucleoSpin Plant II kit (Macherey-Nagel, Germany), a Qubit 4 fluorometer was used for quantitation, and the microarray reaction was performed on an Infinium platform (Illumina, USA) following the protocol of the producer of the Potato V4:0 beadchip kit. The DNA-chips were scanned with an iScan scanner (Illumina, USA), and array data are analysed in GenomStudio.

Results

Results of microarray analysis

The microarray approach in potato research was involved as a high throughput technology for generating saturated linkage maps, which would facilitate the identification of closely linked markers to any mapped phenotype. In the long run, it will promote the breeding of new multi-resistant varieties.

Cultivated potato is a tetraploid plant. Hence the introgression of new characteristics requires a complex approach. Breeding utilisation of molecularly identified valuable genes can be facilitated by applying the microarray technology. The chromosomal distribution of SNP-s on the Potato V4.0 beadchip is revealed in Figure 3.

As revealed in Fig. 3., 2,249 SNP markers can be found on potato chromosome 11, where the PVY extreme resistance gene ***Ry$_{adg}$*** is localised. By analysis of the 65 F1 genotypes of the 89.451 x Somogyi kifli segregating population, it will be possible to identify the chromosomal region where the ***Ry$_{adg}$*** gene resides. With the same SNP-s the chromosomal region can be identified in the White Lady genome, and WL transcripts from this region can be used to identify the homologous 89.451 transcripts. These transcripts will be aligned on the whole genome sequence of 89.451, and that region can be further analysed for resistance gene-like sequences. WGS of 89.451 will be done on our NextSeq 500 sequencer in the coming weeks. Here, we aim to see the chromosomal region where the ***Ry$_{adg}$*** resides.

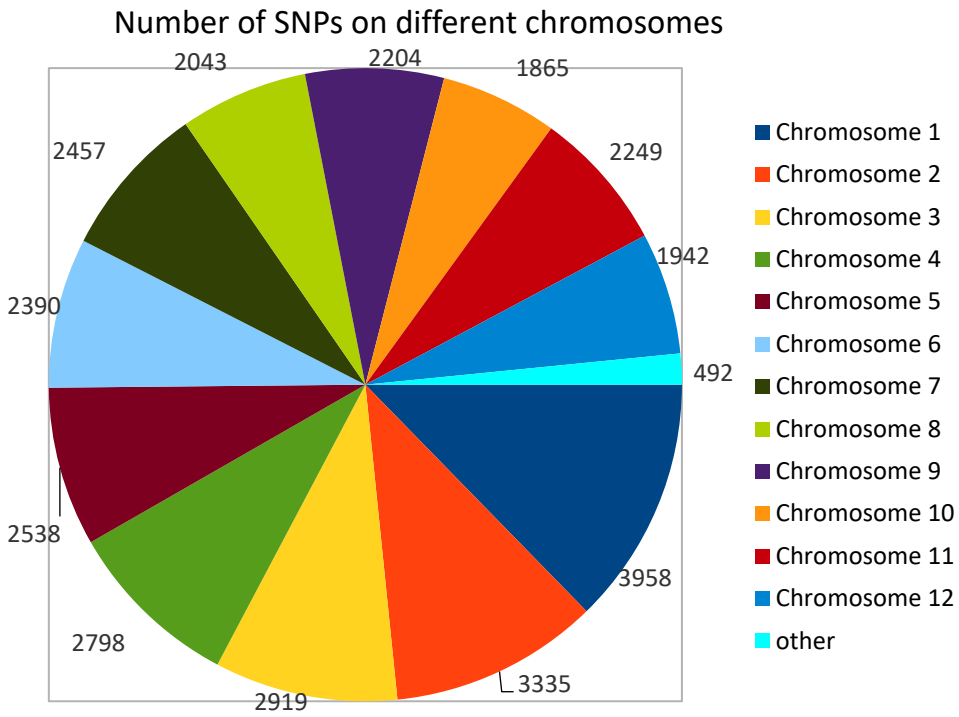

Figure 3. Distribution of SNP-s ont he potato chromosomes

As an example, an SNP graph can be found in Figure 4., showing the distribution of the five possible genotypes created from the SNP alleles.

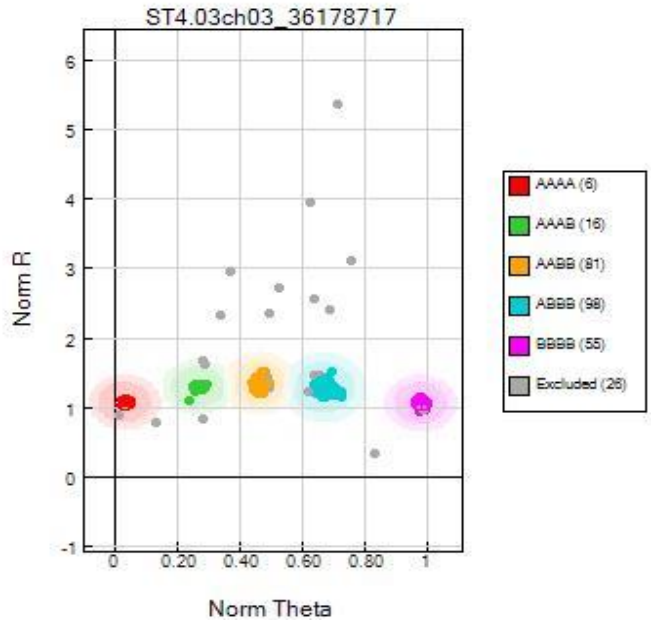

Figure. 4. Distribution graph of an SNP. *NormR means the signal intensity, Norm Theta the allele frequency. In parenthesis, the number of genotypes is revealed.*

Epilogue

One may ask why besides such enormous amount of data, no paper was published in any significant journal. As the principal investigator, I fill that I have briefly to mention some episodes which hindered the development of the project:

In the first year, due to heavy rainfalls, the inland water damaged a significant part of the experimental plant material in the breeding garden of the PRC.

In the second year, the breeder's contract, who was also the director of the PRC was promptly terminated by the university. With him, the determining experts also left. The remaining staff did its best, but it gradually turned out that the experimental results were unreliable, many genotypes were lost, and even the phytophthora isolate could not be adequately maintained.

In the third year, after institutional changes (Univ. of Pannonia, then Szent István Univ., and then MATE), the breeder and the leading Phytopathologist could return and we could restart the project.

From the middle of the last year, we can use a phytotron, we mastered the use of the two HT technologies, and from this year, we could significantly, with five times, expand the server capacity. That means that the bioinformatic analyses will be incomparably more rapid. I'm sure that with these developments, we can fulfil in a reasonable time to explore the targeted results and publish papers from the different segments of this complex project.

Keszthely, March 31. 2023

Dr János Taller
principal investigator