

Comparative Analysis of Alternative Splicing Events in Foliar Transcriptomes of Potato Plants Inoculated with *Phytophthora Infestans*

Jessica A. Lee, Xiangjia Min✉

Department of Chemical and Biological Sciences, Youngstown State University, OH 44555, USA

✉ Corresponding author email: xmin@ysu.edu

Computational Molecular Biology, 2023, Vol.13, No.1 doi: [10.5376/cmb.2023.13.0001](https://doi.org/10.5376/cmb.2023.13.0001)

Received: 08 Oct., 2022

Accepted: 21 Dec., 2022

Published: 29 Jan., 2023

Copyright © 2023 Lee and Min, This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:

Lee J.A., and Min X., 2023, Comparative analysis of alternative splicing events in foliar transcriptomes of potato plants inoculated with *Phytophthora Infestans*, Computational Molecular Biology, 13(1): 1-8 (doi: [10.5376/cmb.2023.13.0001](https://doi.org/10.5376/cmb.2023.13.0001))

Abstract Alternative splicing (AS) is a common process during gene expression of plants in coping various biotic or abiotic stresses. The work reports identification and analysis of AS events in foliar samples of two potato lines, including a wild type line and a pathogen resistant transgenic line (+RB), inoculated with *Phytophthora infestans*. After combining all RNA-seq data collected from 36 samples, a total of 10,246 AS events were identified, including 1,563 exon skipping, 1,368 alternative donor sites, 3,091 alternative acceptor sites, 884 intron retention, and 3,340 complex events, which consisted of more than one basic event. These AS events were generated from 45,874 isoform transcripts expressed from 13,704 genes. It was estimated 30.2% of genes undergoing AS in this analysis. Furthermore, we identified 406 specific AS events, which were generated from 281 genes, and 766 differentially expressed transcripts (DETs) in the sample collected 24 hours after inoculation of *P. infestans* in +RB lines. These DETs were expressed from 763 genes, and among them, 338 genes were alternatively spliced. These results indicate that both AS and differential gene expression may contribute to the resistance against *P. infestans* in +RB line of potato plants.

Keywords Alternative splicing; RNA sequencing; Transcriptome; *Phytophthora infestans*; Blight pathogen

Alternative splicing (AS) is a process of one pre-mRNA transcript generating two or more mRNA isoforms in eukaryotic intron-containing genes. The roles that AS plays in the regulation of growth and development as well as in responding to various stress conditions are well documented in plant species (Staiger and Browns, 2013; Reddy et al., 2013). Our recent analysis and reports from others reveal that ~70% of genes may have pre-mRNAs undergoing AS in plant species (Kim et al., 2007; Chaudhary et al., 2019; Clark et al., 2019).

Potato (*Solanum tuberosum* L.) is one of the most important food crops. A number of genes with pre-mRNA subjected to AS in potato have been reported. For example, under cold stress, an invertase gene generated two transcripts, with one of them having an exon skipping event (Anne-Sophie et al., 1996); BRANCHED1a (BRC1a) gene, which encodes a TCP transcription factor that controls lateral shoot outgrowth, generated two transcripts, with one of them retaining an intron (Nicolas et al., 2015); one of two StPPCK2 (phosphoenolpyruvate carboxylase kinase, PPCK) genes produced an isoform transcript with an intron retained (Marsh et al., 2003); and amongst 8 potato metacaspases (SotubMCs) genes, SotubMC2, SotubMC4, SotubMC6 and SotubMC7 genes produced multiple alternative spliced variants of different lengths (Dubey et al., 2019). The first draft potato genome represented approximately 86% of the 844-megabase genome with 39,031 protein coding genes identified from a homozygous doubled-monoploid potato clone (Potato Genome Sequencing Consortium, 2011). The genome sequence super scaffolds were further assembled into pseudomolecules corresponding to 12 potato chromosomes, representing 674 Mb of the 723 Mb genome assembly having 37,482 protein coding genes (Sharma et al., 2013). RNA sequencing (RNA-seq) is using next-generation sequencing (NGS) technology to sequence and quantify RNA transcripts of a transcriptome (Chu and Corey, 2012). Thus, the available genome sequences and RNA-seq data provide an unprecedented opportunity for genome-wide identification of AS events in potato plants.

Phytophthora infestans is an oomycete, a fungus-like microorganism that causes the serious potato disease known as late blight or potato blight. Gene RB cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight (Song et al., 2003; Van Der Vossen et al., 2003). Using an RNA-seq approach, Gao and Bradeen (2016) collected the foliar transcriptomes in response to *P. infestans* inoculation of two lines of potato plants, including a wild type line and a transgenic line (+RB) which contains the disease resistance (RB) gene against *P. infestans*. Their analysis focused on the identification and analysis of differentially expressed genes (DEGs). Because AS in plants is a common process in development and growth regulation, and in coping various biotic and abiotic stresses, including plant defense against pathogens (Gassmann, 2008), we used RNA-seq data generated by Gao and Bradeen (2016) to examine AS events and differential isoform expressions in foliar transcriptomes of the two contrasting lines of potato plants inoculated with *P. infestans*. This work is part of our efforts aiming at integrating multiple data resources to generate a catalog of genes subjecting to AS in potato plants.

1 Materials and Methods

1.1 Genome and RNA-seq data collection

The genome sequence and annotation files (version 4.03) were downloaded from the Phytozome database (<https://phytozome-next.jgi.doe.gov/>) (Sharma et al., 2013). RNA-seq data from the NCBI Sequence Read Archive (accession number SRP073120) were downloaded (<https://www.ncbi.nlm.nih.gov/sra/docs/srdownload/>) using SRA Toolkit. The dataset consisting of 36 transcriptomes from foliage samples were generated by Gao and Bradeen (2016). The foliar transcriptomes were collected from wild type (WT) and transgenic line (+RB) plants with each inoculated with either *P. infestans* or water and sampled at 0 (pre-inoculation), 6 and 24 hours-post inoculation with three bio-replicates. The details of the plant materials, RNA-preparation and sequencing were described by Gao and Bradeen (2016).

1.2 RNA-seq reads mapping to the genome, analysis of AS events and differentially expressed transcript isoforms (DETs)

The RNA-seq reads were mapped to potato genome sequences using TopHat (v2.2.6) with default parameters (Kim et al., 2013). Then, the transcript alignment file together with genome annotation GFF3 (v4.03) was used as input for Cufflinks (v2.2.1) (<http://cole-trapnell-lab.github.io/cufflinks/>) (Trapnell et al., 2010). The GTF (Gene Transfer Format) files generated from each RNA-seq dataset of bio-replicates after Cufflinks were merged using Cuffmerge script within the Cufflinks package (Trapnell et al., 2010). AStalavista was used for AS event classification (Foissac and Sammeth, 2007). DETs were identified using Cuffdiff script with default settings (FDR adjusted p-value < 0.05) within Cufflinks package (Trapnell et al., 2010).

1.3 Transcripts functional annotation

The transcript sequences were retrieved using the `gtf_to_fasta` tool in the TopHat package (Kim et al., 2013). The annotation information was extracted from the annotation file for previously annotated transcripts (version 4.03), and new transcripts were functionally annotated, using BLASTX search against UniProt-Swiss-Prot database. The GTF file, transcript sequence file, and additional data can be downloaded at: <http://proteomics.yasu.edu/publication/data/Potato/GAO2016/>.

2 Results

2.1 Mapping RNA-seq data to the genome

A total of 36 RNA-seq data were collected from NCBI SRA database which were generated from foliar samples by Gao and Bradeen (2016). The data varied from 7 to 27 million paired reads. These sequence data were mapped to the genome with a mapping rate varying from 29.4 to 77.5% and 1.9 to 5.6% reads mapped to two or more genomic loci, with a total of 547.8 million reads (53.1%) mapped to the reference genome. Among them 3.8% were mapped to multiple loci.

2.2 Identification of AS events and comparative analysis between two lines with or without pathogen treatments

AS events were identified from each of 36 RNA-seq data. Because both the numbers of raw reads and the mapping rates varied substantially among treatments, comparing the numbers of AS events would not provide a meaningful outcome. Intuitively, we expect that the number of AS events would have a strong positive correlation with the number of mapped reads, as more reads mapped to the genome would generate more transcript isoforms and AS events. The correlation analysis of mapped reads and the AS events shows a strong positive linear correlation with a correlation coefficient (R) of 0.97 regardless of the treatments (Figure 1A). That is, 94% (R^2) of the differences among the numbers of AS events in different RNA-seq data can be explained by the numbers of mapped reads. Furthermore, we used cuffmerge tool to merge the mapping gtf files of the three replicates in each treatment for comparing the profile of AS events at different sampling times after inoculation. A trend similar to the individual mapping sample data was found, that is, the AS events in the combined data among treatments were highly positively correlated with the number of mapped reads ($R^2 = 0.95$) (Figure 1B). This was also true when the data collected at different times were further combined, the numbers of mapped reads could explain 98% of the differences in the numbers of AS events in the four treatments ($R^2 = 0.98$).

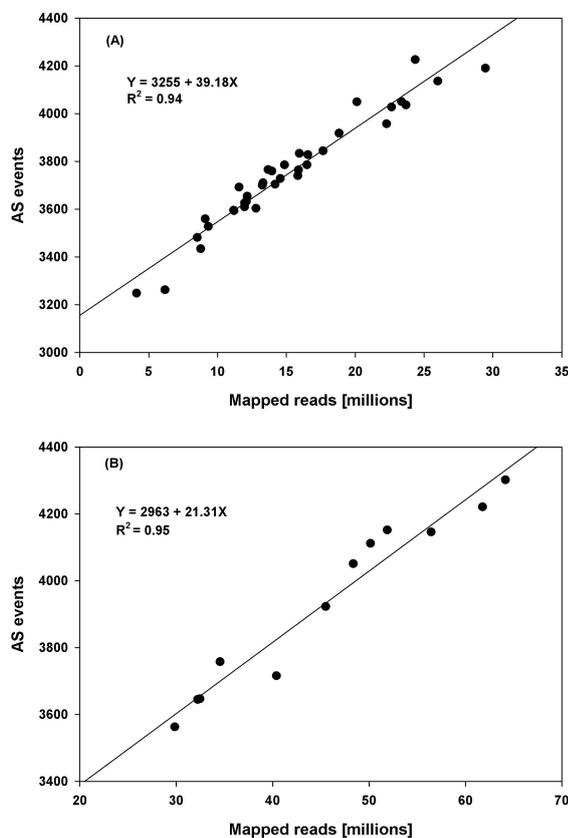


Figure 1 Correlations of mapped reads and the number of alternative splicing events in tomato foliar transcriptome samples
 Note: (A) data from individual transcriptome samples; (B) data from 12 treatments with bio-replicates merged

The categories of AS events in each treatment at different sampling times, merged data for all time points, and all merged data were listed in Table 1. The types of AS events were further classified into four basic types including intron retention (IR), alternative acceptor site (AA), alternative donor site (AD), and exon skipping (ES) (Sablok et al., 2011). Other events or complex events refer to two isoforms having two or more basic AS events. When individual sample data were analyzed, the average proportions were 26.0% AA, 23.8% IR, 15.7% ES, and 13.6% AD, respectively, though there were some variations among different samples. However, when we merged data of sample replicates, the average proportions were 29.9% AA, 18.8% ES, 15.9% AD, and 9.1% IR. The proportion of IR was greatly decreased in comparing with the individual RNA-seq data (Table 1). Similar trends were clearly

observed when further combining data from different sampling times in each treatment (Table 1). We speculate that this is caused by cuffmerge, as this tool filters out a number of lowly expressed “transfrags” that are “probably artifacts.” After the combination of all RNA-seq data generated in this project, a total of 10,246 AS events were identified, including 1,563 ES, 1,368 AD, 3091 AA, 884 IR, and 3,340 complex events (Table 1). These AS events were generated from 13,704 genes with 45,874 isoform transcripts. Since a total of 45,305 genes were identified using RNA-seq mapping, it was estimated 30.2% of genes undergoing AS in potato in this project.

Table 1 Classification of alternative splicing events identified in different treatments of two lines of potato plants

Treatments	Description	Samples	ES	AD	AA	IR	Other	Total
HY0	+RB, Water, 0h	3	753	647	1224	380	1074	4084
HY1	+RB, Water, 6h	3	801	683	1299	405	1114	4302
HY3	+RB, Water, 24h	3	796	650	1299	387	1089	4221
MHY0	+RB, P. inf, 0h	3	752	650	1186	360	1103	4051
MHY1	+RB, P. inf, 6h	3	680	573	1043	307	960	3563
MHY3	+RB, P. inf, 24h	3	697	597	1104	347	1013	3758
MRY0	WT, P. inf, 0h	3	772	663	1220	377	1080	4112
MRY1	WT, P. inf, 6h	3	772	651	1272	383	1074	4152
MRY3	WT, P. inf, 24h	3	754	622	1184	359	1004	3923
RY0	WT, Water, 0h	3	707	567	1077	334	962	3647
RY1	WT, Water, 6h	3	692	571	1083	320	979	3645
RY3	WT, Water, 24h	3	708	611	1097	318	982	3716
HY	+RB, Water	9	1035	925	1956	616	1644	6176
MHY	+RB, P. inf	9	862	781	1539	459	1490	5131
MRY	WT, P. inf	9	1007	847	1741	520	1546	5661
RY	WT, Water	9	865	720	1479	440	1301	4805
WT	All WT	18	1181	960	2095	606	1972	6814
RB	All +RB	18	1201	1093	2390	732	2313	7729
All36	All 36 samples	36	1563	1368	3091	884	3340	10246

Note: ES: exon skipping; AD: alternative donor; AA: alternative acceptor; IR: intron retention; P. inf: *Phytophthora infestans* inoculated

As the number of AS events is greatly dependent on the number of mapped reads, it cannot be used to directly detect the effects of treatments of +RB gene and *P. infestans* inoculation. Thus, we compared the profiles of AS events among treatments to examine the specific changes in AS events of treatment effects. The dynamic changes of AS events were clearly demonstrated in the data sampled at 0 h (pre-inoculation), 6 h and 24 h after inoculation in four different treatments (Figure 2). We carried more analysis on the transcriptomes collected at 24 h post-inoculation and identified 406 specific AS events only occurred in +RB line inoculated *P. infestans* (Figure 2). These AS events involved 899 transcripts generated from 281 genes including 32 newly identified genes in this work. The functions of these genes are diverse and some of them may be involved in the regulation of potato resistance against this pathogen.

3.3 Identification of differentially expressed transcript isoforms (DETs)

Detailed analysis of differentially expressed genes (DEGs) among the treatments were reported by Gao and Bradeen (2016). As there were 30.2% genes alternatively spliced in this study, identifying DETs, in addition to identifying DEGs, was expected to provide a more accurate understanding of the responses to the inoculation of *P. infestans* in the two contrasting potato lines. The numbers of DETs identified between two lines and treatment (water- vs *P. infestans*-inoculation) comparisons at three different sampling time points, using an FDR threshold of 0.05 (Table 2).

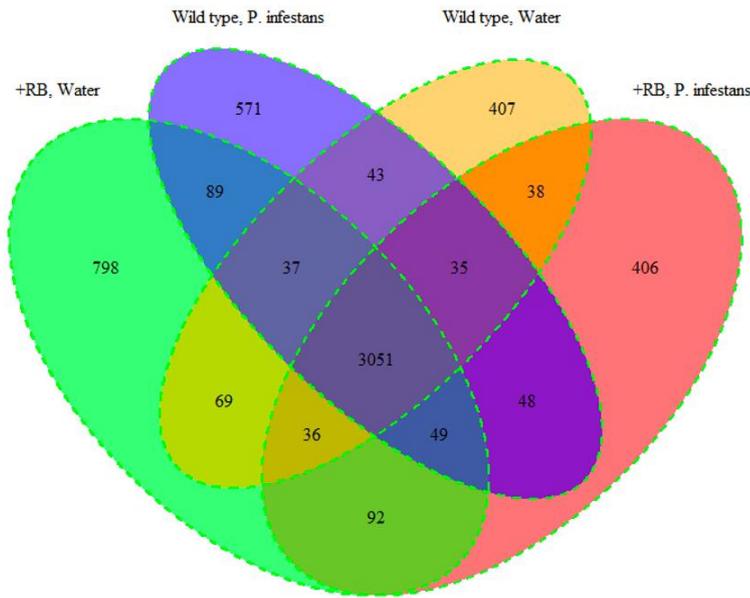


Figure 2 Comparison of alternative splicing events of foliar transcriptomes sampled at 24 hours after inoculation with *P. infestans* and water as controls in wild type and +RB lines of potato plants

Table 2 Identification of differential expressed transcripts in two lines of potato plants pre- or post- inoculation of *P. infestans*

Treatments	0 h	6 h	24 h
+RB (water vs. <i>P. infestans</i>)	325	0	973
WT (water vs. <i>P. infestans</i>)	41	51	182
WT vs. +RB (water)	0	107	110
WT vs. +RB (<i>P. infestans</i>)	537	0	55

At 0 h (pre-inoculation), we would expect there were no DETs in comparison of water and *P. infestans* treatments in either line of potato plants. However, there were 325 DETs identified in +RB line between water treatment and *P. infestans* treatment. In addition, there were 537 DETs identified in comparison of two lines used for *P. infestans* inoculation even pre-inoculation, in contrast, there was no DET identified in these two lines used for water treatments. Such a large discrepancy between these two comparisons was likely caused by sampling errors or unknown factors, although in theory there should be no difference as these samples were collected prior to inoculation of *P. infestans* (Table 2).

Interestingly, at 6 h post-inoculation of *P. infestans* DETs were not detected in the +RB line inoculated with *P. infestans* in comparing with water treatment, however, 973 DETs were detected at 24 h after inoculation in +RB line, while only 182 DETs were detected in wild-type plants for the same treatments (Table 2). Thus, we focused on analyzing the DETs of samples at 24 h and identified 766 DETs specific in the +RB line inoculated with *P. infestans* in comparing water treatment by removing DETs commonly found in other treatments at 24 h samplings (Figure 3). We compared these 766 DETs with those 325 DETs identified at 0 h in +RB line, 732 of them were only found in the 24 h treatments. Further we compared these 766 DETs with 537 DETs identified between two lines pre-inoculation of *P. infestans* (0 h), 700 DETs were only found in the +RB lines at 24 h post-inoculation. Thus, we expected these DETs most likely represented the true +RB line responses to *P. infestans* inoculations. Among these 766 DETs, 347 of them were down-regulated and 419 of them were up-regulated, 505 of them had an adjusted p-value < 0.01. These DETs were generated from 763 genes, and interestingly, 338 of them were alternatively spliced. The functional annotation and their expression values of DETs can be found in the supplementary files available on our website.

References

- Anne-Sophie B., Peter E.H., Anne M., Robbie W., and Gordon C.M., 1996, Exon skipping induced by cold stress in a potato invertase gene transcript, *Nucleic Acids Research*, 24(12): 2347–2351
<https://doi.org/10.1093/nar/24.12.2347>
PMid:8710506 PMCid:PMC145944
- Cao W., Gan L., Wang C., Zhao X., Zhang M., Du J., Zhou S., and Zhu C., 2021, Genome-wide identification and characterization of potato long non-coding RNAs associated with *Phytophthora infestans* resistance, *Front. Plant Sci.*, 12: 619062
<https://doi.org/10.3389/fpls.2021.619062>
PMid:33643350 PMCid:PMC7902931
- Chaudhary S., Khokhar W., Jabre I., Reddy A. S., Byrne L. J., Wilson C. M., and Syed N. H., 2019, Alternative splicing and protein diversity: plants versus animals, *Front. Plant Sci.*, 10: 708
<https://doi.org/10.3389/fpls.2019.00708>
PMid:31244866 PMCid:PMC6581706
- Chu Y., and Corey D.R., 2012, RNA sequencing: platform selection, experimental design, and data interpretation, *Nucleic Acid Therapeutics*, 22(4): 271-274
<https://doi.org/10.1089/nat.2012.0367>
PMid:22830413 PMCid:PMC3426205
- Clark S., Yu F., Gu L., and Min X. J., 2019, Expanding alternative splicing identification by integrating multiple sources of transcription data in tomato, *Front. Plant Sci.*, 10: 689
<https://doi.org/10.3389/fpls.2019.00689>
PMid:31191588 PMCid:PMC6546887
- DeYoung B.J., and Innes R.W., 2006, Plant NBS-LRR proteins in pathogen sensing and host defense, *Nat. Immunol.*, 7(12): 1243-1249
<https://doi.org/10.1038/ni1410>
PMid:17110940 PMCid:PMC1973153
- Dubey N., Trivedi M., Varsani S., Vyas V., Farsodia M., and Singh S.K., 2019, Genome-wide characterization, molecular evolution and expression profiling of the metacaspases in potato (*Solanum tuberosum* L.), *Heliyon*, 5(2): e01162
<https://doi.org/10.1016/j.heliyon.2019.e01162>
PMid:30793051 PMCid:PMC6370574
- Foissac S., and Sammeth M., 2007, ASTALAVISTA: dynamic and flexible analysis of alternative splicing events in custom gene datasets, *Nucleic Acids Res.*, 35: W297–299
<https://doi.org/10.1093/nar/gkm311>
PMid:17485470 PMCid:PMC1933205
- Gao L., and Bradeen J.M., 2016, Contrasting potato foliage and tuber defense mechanisms against the late blight pathogen *Phytophthora infestans*, *PLoS One*, 11(7): e0159969
<https://doi.org/10.1371/journal.pone.0159969>
PMid:27441721 PMCid:PMC4956046
- Gao L., Tu Z.J., Millett B.P., and Bradeen J.M., 2013, Insights into organ-specific pathogen defense responses in plants: RNA-seq analysis of potato tuber-*Phytophthora infestans* interactions, *BMC Genomics*, 14: 340
<https://doi.org/10.1186/1471-2164-14-340>
PMid:23702331 PMCid:PMC3674932
- Gassmann W., 2008, Alternative splicing in plant defense, *Curr Top Microbiol Immunol.*, 326: 219-233
https://doi.org/10.1007/978-3-540-76776-3_12
PMid:18630755
- Kim D., Perlea G., Trapnell C., Pimentel H., Kelley R., and Salzberg S. L., 2013, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions, *Genome Biol.*, 14: R36
<https://doi.org/10.1186/gb-2013-14-4-r36>
PMid:23618408 PMCid:PMC4053844
- Kim E., Magen A., and Ast G., 2007, Different levels of alternative splicing among eukaryotes, *Nucleic Acids Res.*, 35: 125–131
<https://doi.org/10.1093/nar/gkl924>
PMid:17158149 PMCid:PMC1802581
- Marsh J.T., Sullivan S., Hartwell J., and Nimmo H.G., 2003, Structure and expression of phospho enol pyruvate carboxylase kinase genes in Solanaceae. a novel gene exhibits alternative splicing, *Plant Physiol.*, 133(4): 2021-2018
<https://doi.org/10.1104/pp.103.030775>
PMid:14630963 PMCid:PMC300753

- Nicolas M., Rodríguez-Buey M.L., Franco-Zorrilla J.M., and Cubas P., 2015, A recently evolved alternative splice site in the BRANCHED1a gene controls potato plant architecture, *Curr. Biol.*, 25(14): 1799-1809
<https://doi.org/10.1016/j.cub.2015.05.053>
PMid:26119747
- Potato Genome Sequencing Consortium, 2011, Genome sequence and analysis of the tuber crop potato, *Nature*, 475(7355): 189-195
<https://doi.org/10.1038/nature10158>
PMid:21743474
- Reddy A.S., Marquez Y., Kalyna M., and Bartab A., 2013, Complexity of the alternative splicing landscape in plants, *Plant Cell*, 25: 3657-3683
<https://doi.org/10.1105/tpc.113.117523>
PMid:24179125 PMCID:PMC3877793
- Sablok G., Gupta P.K., Baek J.M., Vazquez F., and Min X.J., 2011, Genome-wide survey of alternative splicing in the grass *Brachypodium distachyon*: an emerging model biosystem for plant functional genomics, *Biotechnol. Lett.*, 33(3): 629-636
<https://doi.org/10.1007/s10529-010-0475-6>
PMid:21107652
- Sharma S.K., Bolser D., de Boer J., Sønderkær M., Amoros W., Carboni M. F., D'Ambrosio J. M., de la Cruz G., Di Genova A., Douches D. S., and Eguiluz M., 2013, Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps., *G3* (Bethesda, Md.), 3(11): 2031-2047
<https://doi.org/10.1534/g3.113.007153>
- Song J., Bradeen J.M., Naess S.K., Raasch J.A., Wielgus S.M., Haberlach G.T., Liu J., Kuang H., Austin-Phillips S., Buell C.R., and Helgeson J.P., 2003, Gene RB cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight, *Proceedings of the national academy of sciences*, 100(16): 9128-9133
<https://doi.org/10.1073/pnas.1533501100>
PMid:12872003 PMCID:PMC170883
- Staiger D., and Brown J.W., 2013, Alternative splicing at the intersection of biological timing, development, and stress responses, *Plant Cell*, 25: 3640-3656
<https://doi.org/10.1105/tpc.113.113803>
PMid:24179132 PMCID:PMC3877812
- Trapnell C., Williams B.A., Pertea G., Mortazavi A., Kwan G., Van Baren M. J., Salzberg S. L., Wold B. J., and Pachter L., 2010, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation, *Nat. Biotechnol.*, 28: 511-515
<https://doi.org/10.1038/nbt.1621>
PMid:20436464 PMCID:PMC3146043
- Van Der Vossen E., Sikkema A., Hekkert B.T., Gros J., Stevens P., Muskens M., Wouters D., Pereira A., Stiekema W., and Allefs S., 2003, An ancient R gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato, *Plant J.*, 36(6): 867-882
<https://doi.org/10.1046/j.1365-313X.2003.01934.x>
PMid:14675451