Potato J (2019) 46 (1): 16-23

GENOME WIDE IDENTIFICATION AND DIGITAL EXPRESSION PROFILING OF THE OVATE FAMILY PROTEINS (OFP) IN POTATO

Hemant Balasaheb Kardile[1], Nirmal Kant Sharma[1], Ankita Sharma[2], Rakesh Bhowmick[3], Rahul[1], Monu Kumari[4], Kumar Nishant Choursia[1], Vinay Bhardwaj[1]

ABSTRACT: Ovate family proteins (OFPs) are a class of proteins with diverse physiological functions in plants. These proteins have been identified as the key regulator of fruit shape in tomato. Irregular sized tuber is common problem in potato and severely affect marketable yield. The factors regulating the oval shape of tuber in potato is poorly understood. We are trying to investigate the role of these protein family in regulating the tuber shape in potato. With this objective we have performed the genome wide identification OFPs from the potato genome and performed the digital expression profiling of these genes in tubers as well as fruits. Our *in-sillico* analysis have identified the 24 OFPs with conserved ovate domain. Potato eFP browser generated digital expression profiles reveals that StOFP are expressed in tubers. The underlying mechanism of gene action of StOFPs will become more clearer once the genes are validated using the wet lab experiments.

KEY WORDS: Genome, in-sillico, ovate family proteins (OFPs), potato

INTRODUCTION

The shape and size of horticulture produce (fruit, tuber, cole etc.) often determines the marketable price of the produce. The experimental findings from the last 100 years reveal that the shape of the fruit is controlled by the novel class of plant regulators known as OVATE FAMILY PROTEINS (OFPs) (Hedrick and Booth, 1907; Price and Drinkard, 1908; Liu *et al.,* 2002). These proteins are widely distributed in the plant kingdom and regulate the diverse aspects of growth and development (Wang *et al.,* 2011; Huang *et al.,* 2013). The uniform tuber size and shape is always major concern in potato production and determines the marketable yield. Uniform and consistent tuber size and shape is prime criteria of buyers of potato market chain for purchasing of potatoes either for seed purpose or consumption (Jemison *et al*., 2008).

Tuber shape plays important role during selections in early generations and becomes a desirable factor for physical quality of the tuber(s) and productivity potential in advanced generations. Criteria for evaluation of tuber shape have largely been-qualitative and comparative. Potato utilization largely depends on tuber shape and size where round tubers are mostly preferred for chips and long tubers for French fries. Apart from consumption, tuber dimension information is also useful in designing the harvesting and post harvesting machines (Singh *et al*., 2004). Tubers with distorted shape and non-uniformity can affect the processing and quality of the finished product as well. Indirectly, tuber shape somehow also helps in estimating other quality traits such as tuber density and starch content (Torppa *et al*., 2006).

[1]ICAR-Central Potato Research Institute, Shimla - 171001, Himachal Pradesh, India
[2]Dr YS Parmar University of Horticulture and Forestry, Nauni, Solan - 173230, Himachal Pradesh, India
[3]ICAR-VPKAS, Almora, Uttarakhand, India
[4]DAV Chandigarh, India
Email: kufrihemant@gmail.com

Eight basic shapes have been used by the International Potato Centre gene-bank to describe tuber shapes which include compressed, round, ovoid, obovoid, elliptic, oblong, long oblong and elongated (Lindqvist-Kreuze *et al.,* 2015). However, the most common shapes in commercial potato cultivar are round, oval, long and oblong. Apart from many other factors, number of tubers set per plant often determines the size of tuber. So far only the morphological and machine vision modeling approaches have been exploited to study the tuber shape. Morphological classification plays an important role in characterization of the tuber shape. Using classical genetics approach, a linked QTL marker to_Pt-437059 was identified governing the shape of tuber which is located on chromosome 10 (Lindqvist-Kreuze *et al.,* 2015). Simultaneously, in another report, the shape locus Ro was identified which is also present on chromosome10 and is controlled by multiple alleles in tubers (Chen *et al.,* 2019). This locus was also linked to a marker 1137-CAPSVI, located 0.2 from Ro (Zhu 2015). Apart from chromosome10, several QTLs governing the tuber shape have been identified including loci on chromosome 2 (Sliwka *et al.,* 2008; Prashar *et al.,* 2014), chromosome 5 (Lindqvist-Kreuze *et al.,* 2015), and chromosome 12 (D'hoop *et al.,* 2008). However, Ro locus is considered as the key locus for distinguishing the tuber shape. Still, limited information is available for the factors governing the shape of the tubers. Here in this work we are studying the tissue specific expression of shape governing family of proteins i.e. OFP in potato tubers. With this objective, we have mined the OFP in potato using the orthologous sequences of the tomato and performed the *in-sillico* analysis. The expression patterns of these proteins were studied in fruits and tubers of potato using the digital expression platforms.

MATERIALS AND METHODS

Genome-wide identification and phylogenetic evaluation of OFPs

OFPs were identified by using BLAST searches performed against the potato whole genome sequence annotations provided at the Potato Genome Sequencing Consortium (PGSC) database using the OFP protein sequence from tomato as query. Sequence information for unique top matches were retrieved and further analyzed. Clustal omega was used for homologous sequence alignment using default settings and result was used to construct an un-rooted phylogenetic tree. The maximum likelihood method (MLM) method was applied to construct a phylogenetic tree in which poisson correction, pairwise deletion and bootstrapping served as default values to appraise the reliability of the tree.

Domain identification of StOFP

Conserved domain was identified by executing domain search by Conserved Domains Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) and pfam database (REF) (http://pfam.sanger.ac.uk/). Only significant domain found in protein sequence were considered as a valid domain. Transmembrane domains in protein sequences of the 37 members of the potato MtN3/saliva/SWEET gene family were identified using TMHMM Server v.2.0 (http://www.cbs.dtu.dk/services/TMHMM).

Protein composition of StOFP

To get more information about nature of the StOFP protein, grand average of hydropathy (GRAVY), PI and the molecular weight were predicted by ProtParam tool available on Expert Protein Analysis System (ExPASy) proteomics server (http://www.expasy.ch/tools/protparam.html). The

stability of proteins was determined based on the instability index predicted using the ProtParam (https://web.expasy.org/cgi-bin/protparam/protparam).

Chromosomal location of StOFP genes

The OFP genes were localized on chromosomal map based on their physical position on the 12 *S. tuberosum* genome assembly using MapChart (version2.2) (https://www.wur.nl/en/show/Mapchart.htm) program.

Gene structure analysis StOFP genes

A schematic diagram of the gene structure of StOFP genes consisting of intron/exon structure was constructed using the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/).

Expression profiles of StOFP genes from potato fruit and tubers

Digital expression profiles of StOFPs were generated using the potato eFP browser (http://bar.utoronto.ca/efp_potato/cgi-bin/efpWeb.cgi). Expression profiles in terms of FPKM values were taken for mature and immature fruit and tuber.

RESULTS AND DISCUSSION

Identification of StOFP gene family members, their gene structure and its distribution on chromosomes

A total of 24 OFP genes with confirmed ovate domain were identified in the potato genome and accordingly denoted as StOFP (Table 1). The StOFP genes were named as

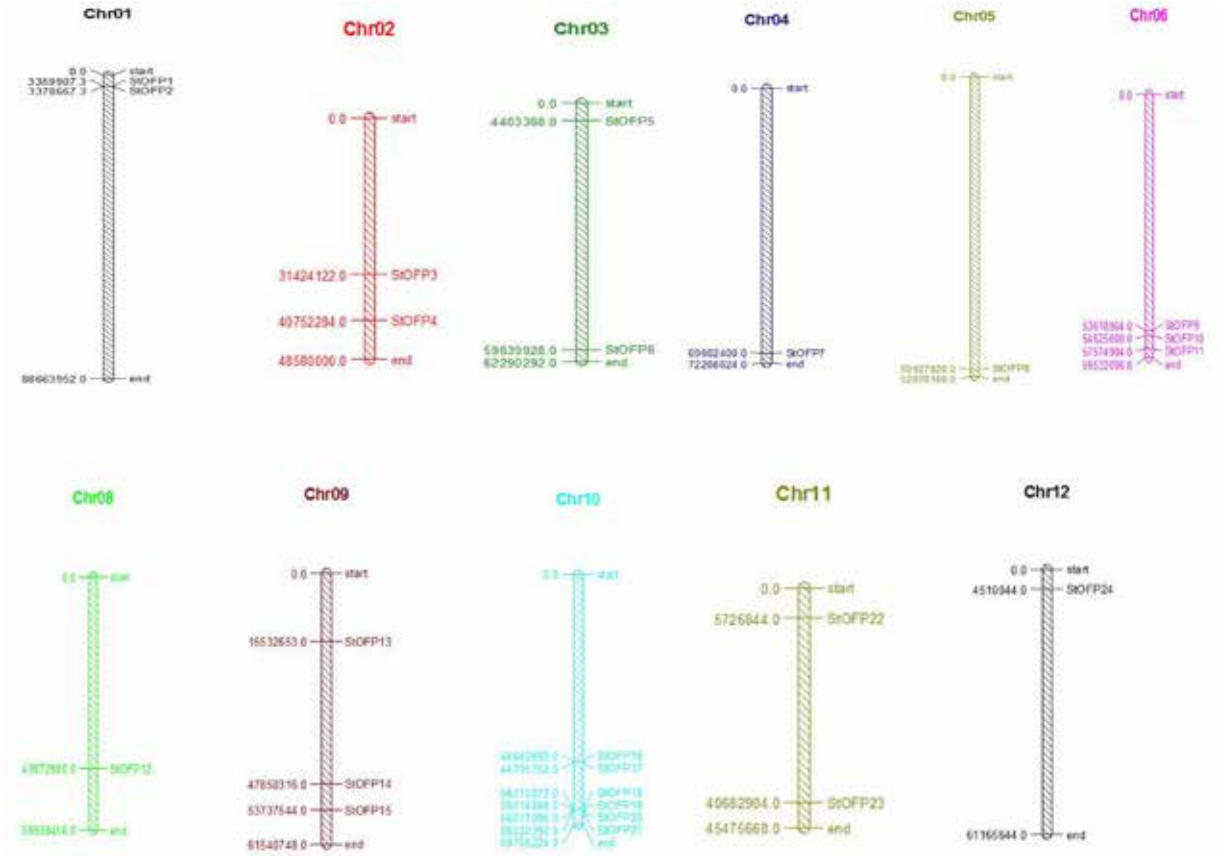

Fig. 1. Chromosomal location showing the different clusters of StOFP genes

Table 1. List of 24 potato OFP genes and their sequence details

Gene name	Chromosome location	Gene ID	mRNA coordinates		Exon	Intron
			Start	End		
StOFP1	ch01	PGSC0003DMG400016415	3359907	3360814	1	0
StOFP2	ch01	PGSC0003DMG400016385	3378667	3379699	1	0
StOFP3	ch02	PGSC0003DMG400028460	31424122	31425454	1	0
StOFP4	ch02	PGSC0003DMG400012688	40752283	40754017	2	1
StOFP5	ch03	PGSC0003DMG400017788	4403368	4405474	3	2
StOFP6	ch03	PGSC0003DMG400005682	59639927	59641275	1	0
StOFP7	ch04	PGSC0003DMG400044629	69602398	69603111	1	0
StOFP8	ch05	PGSC0003DMG400023395	50927918	50928780	1	0
StOFP9	ch06	PGSC0003DMG400027103	53618906	53619587	1	0
StOFP10	ch06	PGSC0003DMG400005900	54625601	54626921	2	1
StOFP11	ch06	PGSC0003DMG400030384	57574904	57576476	1	0
StOFP12	ch08	PGSC0003DMG400014511	43072980	43074248	1	0
StOFP13	ch09	PGSC0003DMG400004418	15532653	15533688	1	0
StOFP14	ch09	PGSC0003DMG400019329	47850315	47851194	1	0
StOFP15	ch09	PGSC0003DMG400032276	53737544	53738921	1	0
StOFP16	ch10	PGSC0003DMG400037114	44642691	44643278	1	0
StOFP17	ch10	PGSC0003DMG400045485	44705754	44706344	1	0
StOFP18	ch10	PGSC0003DMG400040827	56010077	56010829	1	0
StOFP19	ch10	PGSC0003DMG400045504	56014086	56014556	1	0
StOFP20	ch10	PGSC0003DMG400043673	56017095	56018120	1	0
StOFP21	ch10	PGSC0003DMG400028155	56030393	56031156	1	0
StOFP22	ch11	PGSC0003DMG400007601	5726844	5728064	1	0
StOFP23	ch11	PGSC0003DMG400019628	40682903	40683552	1	0
StOFP24	ch12	PGSC0003DMG400000313	4510044	4511292	2	1

per the order of location on chromosomes. The chromosomal location of StOFP gene was determined based on the genomic sequences of potato (**Fig. 1**). A total of three clusters of OFP genes were observed on chromosome 6 and 9. Chromosome 10 harbors highest number (six) of StOFP genes. Apart from these gene clusters rest all OFP genes found to be uniformly distributed on different chromosomes.

Predicted gene structure reveals that (**Fig. 2**) StOFP5 contains three exons, maximum in number followed by two exons in StOFP4, 10, and 24. Interestingly, Twenty out of 24 StOFP are intron less and four genes mainly StOFP4, StOFP5, StOFP10 and StOFP24 contain the intronic regions in their structure.

In-sillico analysis of StOFP proteins

In-sillico analysis was performed for the amino acid sequences of the 24 StOFP proteins. The length of StOFP proteins is ranging from 95 (StOFP9) to 422 (StOFP12) amino acids with lowest molecular weight 11.09 kDa and highest of 48.05 kDa respectively (**Table 2**). Out of 24 StOFP only two proteins StOFP6 and StOFP9 were

Table 2. Number of amino acids, GRAVY and instability index

Gene name	Protein	Mw (KDa)	Amino acid	GRAVY	The instability index (II)	Remark
StOFP1	PGSC0003DMP400028716	20.58	184	-0.532	58.18	unstable
StOFP2	PGSC0003DMP400028626	33.31	302	-0.322	44.57	unstable
StOFP3	PGSC0003DMP400049548	48.18	410	-1.034	56.59	unstable
StOFP4	PGSC0003DMP400022490	40.4	350	-0.818	68.33	unstable
StOFP5	PGSC0003DMP400031077	45.69	397	-0.827	54.02	unstable
StOFP6	PGSC0003DMP400010028	32.25	288	-0.735	36.67	stable
StOFP7	PGSC0003DMP400066733	27.39	237	-0.773	64.3	unstable
StOFP8	PGSC0003DMP400040453	26.16	233	-0.492	49.97	unstable
StOFP9	PGSC0003DMP400047088	11.09	95	-0.266	22.13	stable
StOFP10	PGSC0003DMP400010449	24.43	212	-0.337	52.37	unstable
StOFP11	PGSC0003DMP400052892	39.9	351	-0.862	66.73	unstable
StOFP12	PGSC0003DMP400025572	48.05	422	-0.903	44.65	unstable
StOFP13	PGSC0003DMP400007843	32.51	284	-0.743	67.72	unstable
StOFP14	PGSC0003DMP400033582	25.57	228	-0.621	54.01	unstable
StOFP15	PGSC0003DMP400055486	24.95	220	-0.508	56.93	unstable
StOFP16	PGSC0003DMP400059218	21.5	195	-0.579	62.08	unstable
StOFP17	PGSC0003DMP400067589	21.75	196	-0.511	61.49	unstable
StOFP18	PGSC0003DMP400062931	28.35	250	-0.43	54.34	unstable
StOFP19	PGSC0003DMP400067608	17.42	156	-0.318	53.5	unstable
StOFP20	PGSC0003DMP400065777	38.34	341	-0.474	57.39	unstable
StOFP21	PGSC0003DMP400048910	20.71	189	-0.412	64.76	unstable
StOFP22	PGSC0003DMP400013451	26.62	236	-0.636	57.74	unstable
StOFP23	PGSC0003DMP400034100	15.45	139	-0.636	54.33	unstable
StOFP24	PGSC0003DMP400000623	32.44	280	-0.932	55.65	unstable

stable having instability index 36.67 and 22.13 respectively. All other proteins were unstable having instability index more than 40. All proteins having GRAVY value with maximum -0.266 for StOFP9 and -1.034 for StOFP3. Conserved domain analysis revealed that OFP protein consists of ovate domain as their signature domain.

Digital expression profiling of StOFPs in fruits and tubers

Digital expression profile was generated for the berries as well as the tubers. Expression profiles from the Immature and mature berries revealed that immature berries showed the expression of total 8 genes namely StOFP1, 2, 4, 6, 11, 12, 22 and 23 whereas number of expressed genes is reduced to half in mature berries (**Fig. 3**). Out of these eight isoforms, the four isoforms mainly, StOFP1, 2, 12 and 23 become repressible in mature berries indicating their role in early period of berries development.

Expression profiles from the tuber samples revealed that three isoforms mainly, StOFP 4, 11 and 12 expressed in tubers **(Fig. 4)**. Interestingly the expressed isoforms are same as they are expressed in the berries. Several QTLs for the tuber shape have been

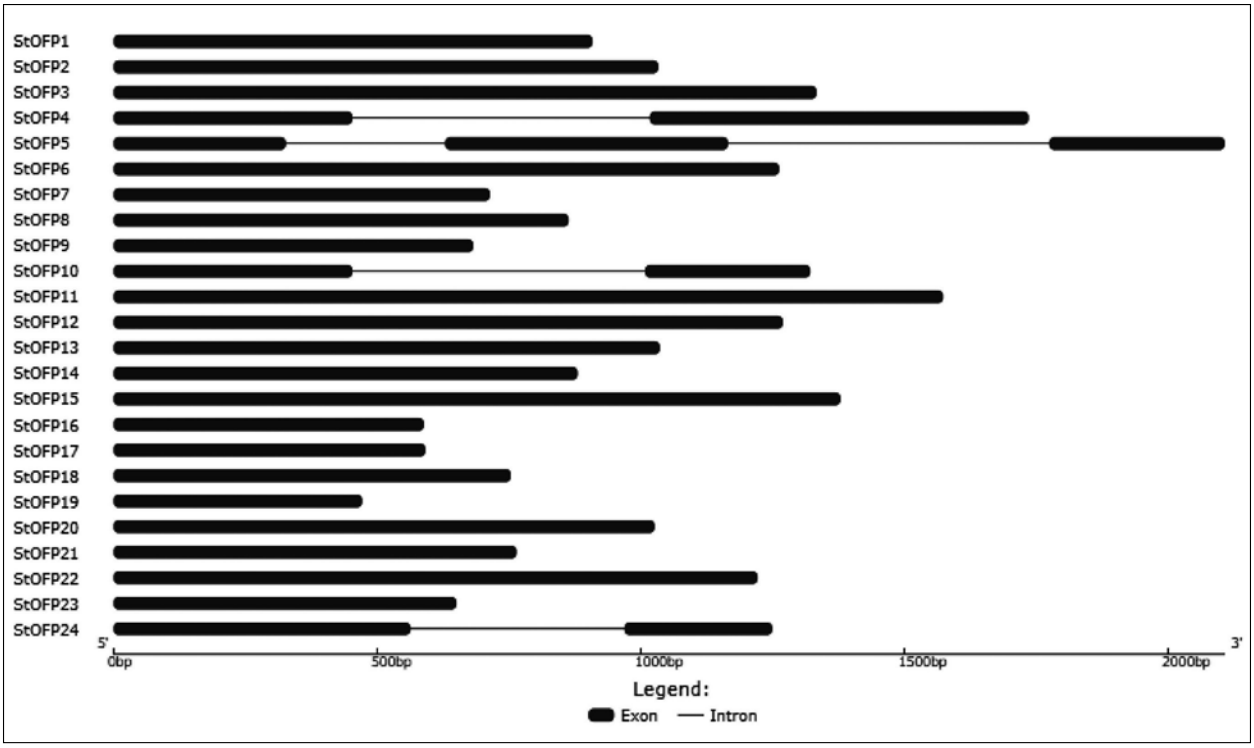

Fig. 2. Gene structure showing exon and intron boundaries of StOFPs.

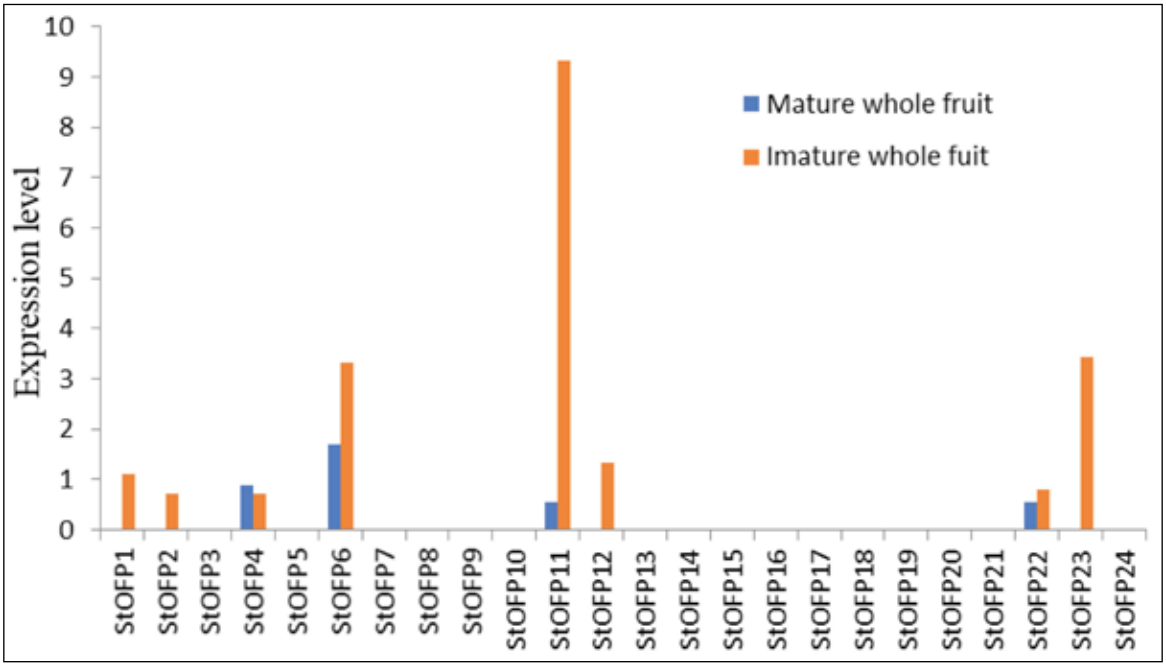

Fig. 3. Digital expression profile of StOFP genes from immature and mature. berries

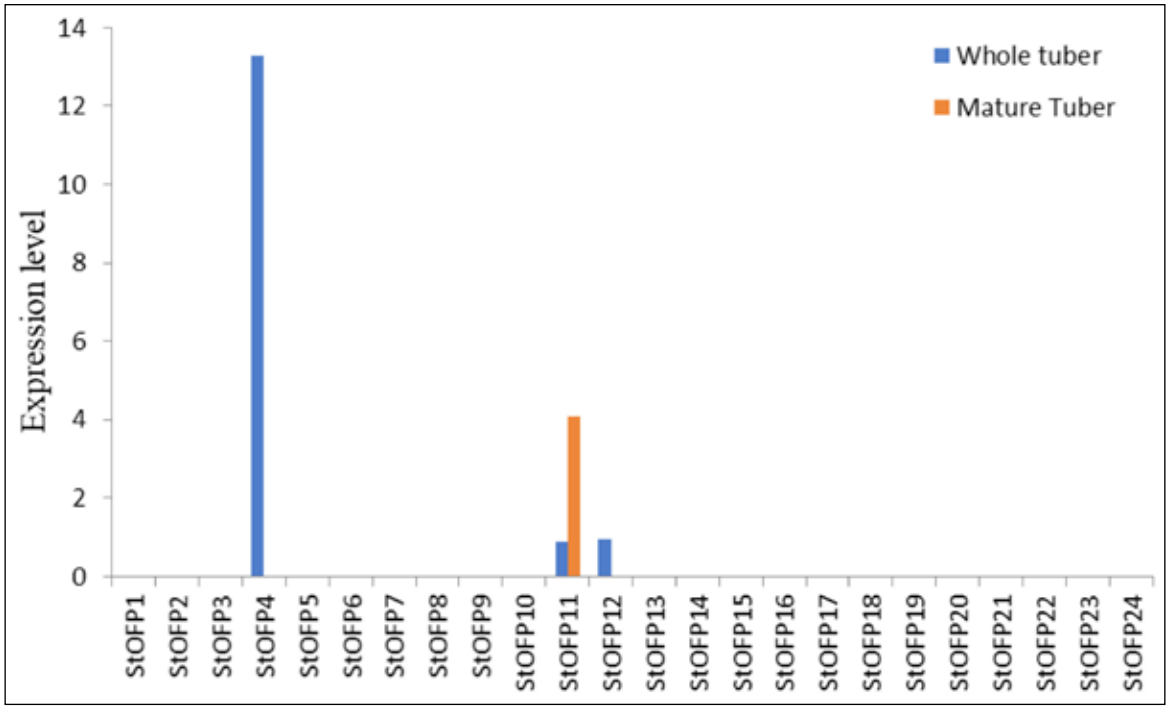

Fig. 4. Digital expression profile of StOFP genes from whole and mature tubers.

mapped on the chromosome 2, 5 and 10 (Sliwka *et al.*, 2008; Prashar *et al.*, 2014) and interestingly one of the isoforms (StOFP 4) expressed in tuber tissues is also located on the chromosome 2. StOFPs expressed in the tubers can be potential targets for the gene manipulation to identify its role in determining the shape of the tubers. The other isoforms which have expressed in the berries may have role in determining the oval shape of the potato berries, but we are not targeting these genes for future experiments as berries has no commercial value.

CONCLUSIONS

Digital expression analysis of StOFPs revealed that three StOFP genes 4, 11, 12 are expressed in the tubers and may have role in determining the shape of the tubers. These genes may be selected for the future experiments to get the deep insight of their role in determining the shape of the tubers.

LITERATURE CITED

Hedrick UP, Booth NO (1907) Mendelian characters in tomato. *Proc Am Soc Hort Sci* **5**: 19-24

Price HC, Drinkard AW (1908) Inheritance in tomato hybrids. *Va Agric Exp Stn Bull* **177**: 17–53.

Liu J, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc Natl Acad Sci USA* **99**: 13302-6.

Wang S, Chang Y, Guo J, Zeng Q, Ellis BE, Chen JG (2011) Arabidopsis ovate family proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development. *PLoS One* **6(8)**: e 23896.

Huang Z, Van Houten J, Gonzalez G, Xiao H, van der Knaap E (2013) Genome-wide identification, phylogeny and expression analysis of SUN, OFP and YABBY gene family in tomato. *Mol Genet Genomics* **288(3-4)**: 111-29

Singh S, Kumar D and Singh BP (2004) Quantification of tuber shape in Indian potato cultivars. *Potato J* **31**: 3-4

Torppa J, Valkonen J and Muinonen K (2006) Three-

dimensional Stochastic Shape Modelling for Potato Tuber. *Potato Research* **49:** 109–118

Lindqvist-Kreuze1 H, Khan A, Salas E, Meiyalaghan S, Thomson S, Gomez R and Bonierbale M (2015) Tuber shape and eye depth variation in a diploid family of Andean potatoes. *BMC Genetics* **16:** 57

Chen N, Zhu W , Xu J, Duan S, Bian C, Hu J, Wang W, Li G and Jin L (2019) Molecular marker development and primary physical map construction for the tuber shape Ro gene locus in diploid potato (*Solanum tuberosum* L.) *Mol Breeding* **39**: 6

Zhu WW (2015) Genetic mapping and molecular markers development of tuber shape gene in potato. Master thesis, Chinese Academy of Agricultural Sciences.

Sliwka J, Wasilewicz-Flis I, Jakuczun H, Gebhardt C (2008) Tagging quantitative trait loci for dormancy, tuber shape, regularity of tuber shape, eye depth and flesh color in diploid potato originated from six Solanum species. *Plant Breed* **127**: 49–55

Prashar A, Hornyik C, Young V, McLean K, Sharma SK, Dale MF, Bryan GJ (2014) Construction of a dense SNP map of a highly heterozygous diploid potato population and QTL analysis of tuber shape and eye depth. *Theor Appl Genet* **127**: 2159–71

Lindqvist-Kreuze H, Khan A, Salas E, Meiyalaghan S, Thomson S, Gomez R, Bonierbale M (2015) Tuber shape and eye depth variation in a diploid family of Andean potatoes. *BMC Genet* **16**: 57

Jemison Jr JM, Sexton P, Camire ME (2008) Factors Influencing Consumer Preference of Fresh Potato Varieties in Maine. *Am J Pot Res* **85**: 140

MS received : 29 May 2019; Accepted: 29 July 2019