Biochemical Genetics (2019) 57:747–766
https://doi.org/10.1007/s10528-019-09920-6

ORIGINAL ARTICLE

Expanding *Phaseolus coccineus* Genomic Resources: De Novo Transcriptome Assembly and Analysis of Landraces 'Gigantes' and 'Elephantes' Reveals Rich Functional Variation

Aliki Xanthopoulou, et al. ***[full author details at the end of the article]***

Received: 20 August 2018 / Accepted: 1 April 2019 / Published online: 17 April 2019

Abstract

Beans are one of the most important staple crops in the world. Runner bean (*Phaseolus coccineus* L.) is a small-scale agriculture crop compared to common bean (*Phaseolusvulgaris*). Beans have been introduced to Europe from the Central America to Europe and since then they have been scattered to different geographical regions. This has resulted in the generation of numerous local cultivars and landraces with distinguished characters and adaptive potential. To identify and characterize the underlying genomic variation of two very closely related runner bean cultivars, we performed RNA-Seq with de novo transcriptome assembly in two landraces of *P. coccineus*, 'Gigantes' and 'Elephantes' phenotypically distinct, differing in seed size and shape. The cleaned reads generated 37,379 and 37,774 transcripts for 'Gigantes' and 'Elephantes,' respectively. A total of 1896 DEGs were identified between the two cultivars, 1248 upregulated in 'Elephantes' and 648 upregulated in 'Gigantes.' A significant upregulation of defense-related genes was observed in 'Elephantes,' among those, numerous members of the *AP2-EREBP*, *WRKY*, *NAC*, and *bHLH* transcription factor families. In total, 3956 and 4322 SSRs were identified in 'Gigantes' and 'Elephantes,' respectively. Trinucleotide repeats were the most dominant repeat motif, accounting for 41.9% in 'Gigantes' and 40.1% in 'Elephantes' of the SSRs identified, followed by dinucleotide repeats (29.1% in both cultivars). Additionally, 19,281 putative SNPs were identified, among those 3161 were non-synonymous, thus having potential functional implications. High-confidence non-synonymous SNPs were successfully validated with an HRM assay, which can be directly adopted for *P. coccineus* molecular breeding. These results significantly expand the number of polymorphic markers within *P. coccineus* genus, enabling the robust identification of runner bean cultivars, the construction of high-resolution genetic maps, potentiating genome-wide association studies. They finally contribute

Electronic supplementary material The online version of this article (doi:https://doi.org/10.1007/s10528-019-09920-6) contains supplementary material, which is available to authorized users.

to the genetic reservoir for the improvement of the closely related and intercrossable *Phaseolus vulgaris*.

Keywords *Phaseolus coccineus* · RNA-seq · Comparative transcriptomics · Landraces · SNPs

Introduction

The genus *Phaseolus* spp. comprises of species valuable for human nutrition, producing seeds rich in protein and essential nutrients (Vaz Patto et al. 2015). Their ability to fix and convert atmospheric nitrogen to organic compounds enhances agricultural sustainability in intercropping and crop rotation systems, reducing fertilizers use (Rubiales and Mikic 2015). The species *Phaseolus vulgaris* (common bean, $2n=2x=22$), domesticated in S. America, is the most widely cultivated worldwide. Due to its economic importance, a plethora of genomics resources have been developed for *P. vulgaris*, such as sequence-based SNP markers (Hyten et al. 2010; Schröder et al. 2016), reference genome sequences for both Andean and Mesoamerican originated cultivars (Vlasova et al. 2016; Schmutz et al. 2014), and database repositories (Li et al. 2016; Bhawna et al. 2016).

Phaseolus coccineus (scarlet runner bean, $2n=2x=22$) is the third most economically important *Phaseolus* species (following *P. lunatus)* (Schwember et al. 2017). It originates probably from Central America, was introduced to Europe along with *P. vulgaris* and spread across different latitudes due to its enhanced adaptation to cooler temperatures and the shorter growing period, compared to *P. vulgaris* (Rodriguez et al. 2013). Contrary to self-pollinated *P. vulgaris*, it is mainly a cross-pollinated species with a low occurrence of selfing as indicated by the higher heterozygosity index (H_o: 0.12–0.31) compared to *P. vulgaris* ($H_o=0.008$–0.10) (Schwember et al. 2017).

Yet, *P. coccineus* is intercrossable with *P. vulgaris.* Interspecies hybridization produces viable seeds especially when the latter is used as the female parent (Escalante et al. 1994; Rodriguez et al. 2013). Natural hybridization also occurs when the two species are grown in proximity (Acosta-Gallegos et al. 2007). The introduction of *P. coccineus* into Europe has resulted in a moderate genetic bottleneck for this species (Rodriguez et al. 2013; Bitocchi et al. 2017). Taking into account the relatively high genetic variability that has been reported (Rodriguez et al. 2013), runner bean can serve as a valuable donor species for inter-specific hybridization in *Phaseolus* breeding programs. In particular, it can serve as a source of diversity for cold tolerance (Rodiño et al. 2007), increased number of pods per inflorescence (Schwember et al. 2017), and resistance to foliar and soil-borne fungi (Schwartz et al. 2006; Schwember et al. 2017).

However, there are scarce data on *P. coccineus* genomics, with few molecular studies and few molecular tools developed, which hampers targeted breeding efforts to increase the utilization of this species. Nuclear and chloroplast simple sequence repeats (SSRs) markers have been developed and deployed for the

characterization of genetic diversity among *P. coccineus* accessions and the estimation of genetic bottlenecks, by the comparison of European and Mesoamerican accessions (Acampora et al. 2007; Sicard et al. 2005; Rodriguez et al. 2013; Spataro et al. 2011). Recently, as a part of a large sequence project of different *Phaseolus* species, the genome of *P. coccineus* was sequenced, indicating significant values of genetic divergence with *P. vulgaris*, but also uncovering gene flow towards *Phaseolus vulgaris*, with an over-representation of loci controlling pathogen defense traits (Rendón-Anaya et al. 2017).

Transcriptome sequencing is a valuable method which provides genomic resources in non-model species with limited genetic resources. It offers valuable information regarding genic polymorphisms that can be functionally important as they can cause potential changes in the proteins expressed. The genetic polymorphisms can be assessed by molecular markers such as SSRs and SNPs (Xanthopoulou et al. 2017) and can be further used for the construction genetic maps and marker-assisted selection. Since genic-SSRs have higher rate of transferability across related species than anonymous genomic SSRs (Xanthopoulou et al. 2014), they are preferable for phylogenetic studies and comparative mapping to pave the way to broader genomic studies.

Next-generation sequencing systems have enabled the generation of large molecular marker datasets including SSR and SNP markers; the discovery of novel genes; the study of gene families, species phylogenetics, and evolution; and the analysis of the transcriptome and its association with physiological and metabolic pathways (Kumar et al. 2012; Torre et al. 2016). The identification of SSRs or SNPs within or near the coding sequences can directly affect the phenotype. Thus, these markers are exceptionally useful in crop improvement programs and quantitative trait locus (QTL) and genome-wide association analyses (GWAS) (Kage et al. 2016).

The Greek landraces of *P. coccineus*, 'Elephantes' ('E') and 'Gigantes' (G'), significantly differing in seed size (Suppl. Figure 1), are traditionally grown as 'dry beans' in the area of Prespa (NW Greece). In this study, we performed RNA-Seq with de novo transcriptome assembly in these two landraces. Significant variation between the two landraces was uncovered, with 19,281 putative SNPs identified. Among those, 3161 were non-synonymous, causing amino acid substitutions and thus might have potential functional implications. A significant upregulation of defense-related genes was observed in 'E.' These results can greatly contribute in resources to discover novel genes and polymorphisms with the aim to tag phenotypic variation in valuable agronomic traits in runner bean with genetic polymorphism that can be used as molecular markers in breeding programs. Additionally, they can contribute to the molecular identification and characterization of runner bean cultivars and landraces, enabling the traceability and authentication of economically important local landraces and/or varieties with the potential to obtain a Protected Designation of Origin (PDO) label. Finally, they provide a reservoir of genes and polymorphisms for common bean improvement through marker-assisted selection.

Materials and Methods

Plant Material

The runner bean landraces 'Gigantes' ('G') and 'Elephantes' ('E'), designated as Protected Geographic Indication (PGI) products, were grown in Prespes (40°50′09.3″N, 21°08′24.6″E), Greece, under the same standard field conditions, using standard horticultural procedures. Leaves were harvested from five plants of each landrace at mature stage and immediately frozen in liquid nitrogen for further experiments.

Sequencing

Total RNA was extracted from each tissue using the TRIzol® Reagent (Invitrogen, USA) and DNaseI (Invitrogen) was then used to remove DNA. The integrity and purity of the total RNA were evaluated using 2% denaturing gel electrophoresis and NanoDrop DU8000 spectrophotometry (A260/A280 and A260/A230). Libraries were prepared from total RNA using a TruSeq stranded total RNA sample preparation kit, the Ribo-Zero plant kit (Catalog No. RS-122–2401; Illumina). Sequencing was performed by Macrogen Inc., South Korea, with HiSeq 2500 using a TruSeq SBS kit version 4 (Illumina).

Assembly and Differential Expression Analysis

Raw reads were pre-processed to remove low-quality regions and adapter sequences, using Trim Galore! (https://cutadapt.readthedocs.org/en/stable/) (version 0.4.0), a wrapper tool around Cutadapt (https://cutadapt.readthedocs.org/en/stable/) and FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to consistently apply quality and adapter trimming to FastQ files. The refined sequence data were used for the de novo transcriptome assembly of the two cultivars, through the Trinity tool as previously described (Grabherr et al. 2011). De novo variant identification and quantification was performed using KisSplice (https://kissplice.prabi.fr). Microsatellite identification software (MISA, https://pgrc.ipkgatersleben.de/misa/) was used to identify SSR markers.

Read abundance was estimated per sample using RSEM by utilizing the corresponding tools provided in the Trinity platform. The expected number of Fragments per Kilobase of transcript sequence per millions base pairs sequenced (FPKM) of each gene was calculated based on the length of the gene and reads count mapped to this gene. Differential expression analysis of the two landraces (five biological replicates each) was performed using the DESeq2 tool which is embedded in Trinity software. Genes with absolute log2 fold change > 2 and an adjusted P value < 0.05 found by DESeq2 were considered as differentially expressed.

Annotation

TransDecoder 2.0.1 (https://github.com/TransDecoder/TransDecoder/wiki) was used to identify candidate coding regions within the generated transcriptome and looked for an open reading frame of at least 100 amino acids long. Trinotate (https://trinotate.github.io/) was used to perform the functional annotation of the transcriptome using the transdecoder-generated peptide sequence file and BLASTX v2.3.1 and BLASTP v2.3.0+ searches with an *E* value cutoff set to 10^{-6}. Functional domains were identified using the Pfam domain database using HMMER 3.0.. Gene Ontology (GO) terms were summarized and visualized using REVIGO (https://revigo.irb.hr/). KEGG pathways were mapped to annotated organisms of the Pea family (namely, *Glycine max*, *Arachis duranensis*, *Arachis ipaensis,* and *Lotus japonicas*) using the KEGG Automatic Annotation Server (https://www.genome.jp/tools/kaas/). KEGG orthology (KO) assignments were performed based on the bi-directional best hit (BHH) of BLAST. GOslim instances were realized using the CateGOrizer tool (https://www.animalgenome.org/cgi-bin/util/gotreei).

Genotyping of SNPs and SSRs Using High-Resolution Melt Assay (HRM)

For the SNP validation, primers suitable for HRM analysis were designed from each landrace and evaluated in 40 individuals (20 'E' and 20 'G') by PCR amplification, DNA melting, and endpoint fluorescence level acquiring. PCR amplifications were performed in a total volume of 15 µL on a Rotor-Gene 6000 real-time 5P HRM PCR Thermocycler (Corbett Research, Sydney, Australia) according to Ganopoulos et al. (2013). The sequences of the SNP primers used in the study are shown in Supplementary Table 1. HRM was performed as described previously by Ganopoulos et al. (2012). For SSR analysis, three primers (Supplementary Table 2) have been employed in the same 40 individuals following Ganopoulos et al. (2012).

Results

Transcriptome Sequencing and Annotation of *P. coccineus* Landraces 'Gigantes' and 'Elephantes'

The transcriptome of the runner bean (*P. coccineus*) landraces, 'G' and E,' with two morphological distinct types characteristically differing in seed size and morphology (Table 1; Supplementary Figure 1) was analyzed. After removing duplicate reads, trimming adaptors, and low-quality reads, a total of 15,952,588 and 16,389,736 reads were obtained for 'G' and 'E,' respectively, which have been deposited in the NCBI Sequence Read Archive (accession number SRP150502). The filtered and trimmed reads were assembled using the Trinity software (Grabherr et al. 2011), generating 37,379 and 37,774 transcripts for 'G' and 'E,' respectively (Table 1).

Table 1 Summary statistics of assembled transcripts

Items	Gigantes	Elephantes
Number of sequences	37,379	37,774
Total length	28,998,460	29,869,930
Mean length	776	791
N50	1203	1224
Maximal length	9045	7011

The total lengths of the assembled transcripts for 'G' and 'E' were 28,998,460 bp and 29,869,930 bp, respectively, with an average length of 776 and 779.

The annotation results for all transcripts in 'G' and 'E' are included in the Supplementary Table 3. BLASTX hits were used for transcript mapping and the subsequent assignment of gene ontology (GO) annotations. Transcripts with at least one mapped GO term were assigned to functional groups belonging to the GO ontologies cellular component, molecular function, and biological process. Overall, the majority of transcripts assigned to GO terms were related to biological processes (32.03% for 'G' and 32.05% for 'E') (Suppl. Figure 2). The top GO terms and their ranking were similar between 'G' and 'E.' In particular, biological processes such as "metabolism," "biosynthesis," "nucleotide and nucleic acid metabolism," "cell organization and biogenesis," and "transport" were among the top GO terms in both landraces. Molecular functions have been assigned to 36.01% and 36.07% of ontologies for 'G' and 'E,' respectively, with the most represented categories being related to "catalytic activity," "transferase activity," "hydrolase activity," and "binding." The remaining 24.93% and 25.23% of annotated transcripts for 'G' and 'E,' respectively, have shown cellular component GO term, mostly related to "intracellular," "cytoplasm," and "nucleus."

Functional Classification of DEGs Between the Two *Phaseolus coccineus* Landraces

Normalized-FPKM (fragments per kilobase per million) were used to quantify the transcript levels. DEGs were defined as genes that were significantly upregulated in one landrace relative to the other one. Between 'E' and 'G,' a total of 1896 DEGs were identified, including 1248 upregulated in 'E' and 648 upregulated in 'G' (Supplementary Table 4).

When assessing DEGs upregulated in 'G,' the top GO terms were "response to stimulus," "response to abiotic stimulus" for metabolic processes (Fig. 1a), "cytoplasmic part" and "membrane" for cellular component (Fig. 1b), and "catalytic activity" and "ion binding" for molecular function (Fig. 1c). Similarly, the highest-ranking categories of biological processes of genes upregulated in 'E' were "response to stimulus" and "response to stress" (Fig. 1d). By contrast, the most represented categories of cellular component were related to membranes (Fig. 1e), and of molecular functions to transcription factor activity (Fig. 1f).

Overrepresented GO terms formed by DEGs between the two landraces were clustered in groups such as "response to biotic stimulus" and "defense response,"

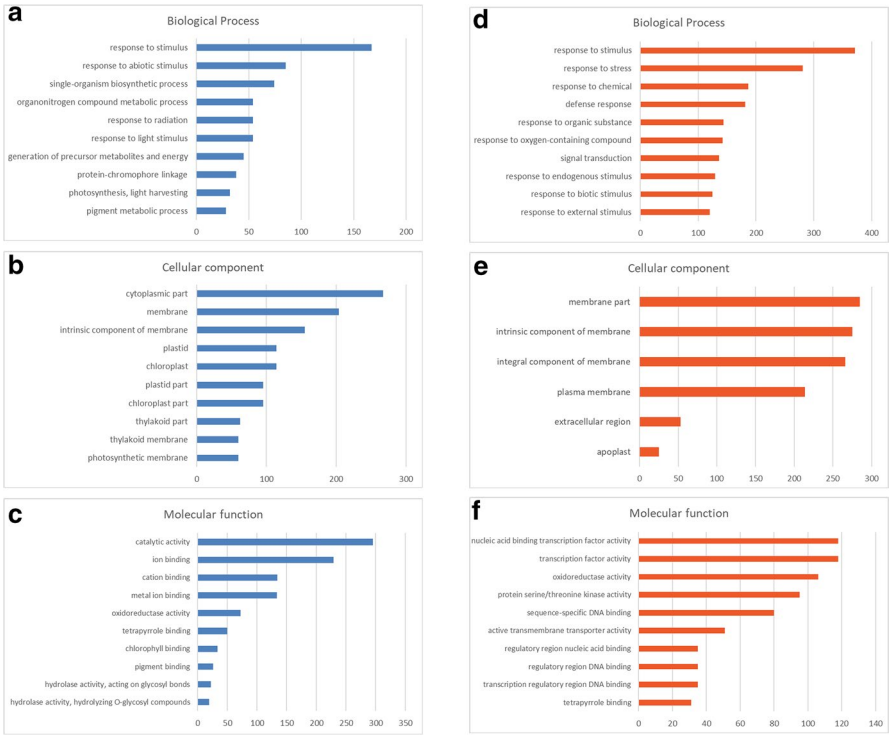

Fig. 1 GO classification by biological process, cellular component, and molecular function of DEGs upregulated in **a**–**c** 'Gigantes' and **d**–**f** 'Elephantes' using GOslim instances realized with CateGOrizer tool. DEGs were considered those with an absolute log2 fold change > 2 and an adjusted *P* value (P_{adj}) < 0.05 using DESeq2 package

"tetrapyrrole metabolic process" and "organonitrogen compound catabolic process," as well as "secondary metabolic process," "flavonoid biosynthetic process," and "pigment biosynthetic process" (Fig. 2). When assessing DEGs upregulated in 'G,' a clustering of the GO terms "response to abiotic stimulus," "response to toxic substance," and "biogenesis" and "pigment biosynthetic" and "metabolic process" was observed (Suppl. Figure 3a). On the other hand, regarding GO terms upregulated in 'E,' clusters were formed related to immune response such as response to biotic stimulus and to chitin, defense response, and secondary metabolic process, as well as metabolite synthesis and transport, such as monocarboxylic acid biosynthesis and transport, S-adenosyl-methionine biosynthesis, and $_{L}$-phenylalanine catabolic process (Suppl. Figure 3b).

When examining the KEGG Pathways for enrichment of genes upregulated in either 'G' or 'E' landraces, a significant upregulation of beta-glucosidases, a sucrose-phosphate synthase, and 1,4-alpha-glucan branching enzyme was observed for 'G' within the starch and sucrose metabolism pathway (Suppl. Figure 4). On the other hand, a striking over-representation of 'E'-upregulated DEGs was observed in plant pathogen interaction, plant hormone signal transduction,

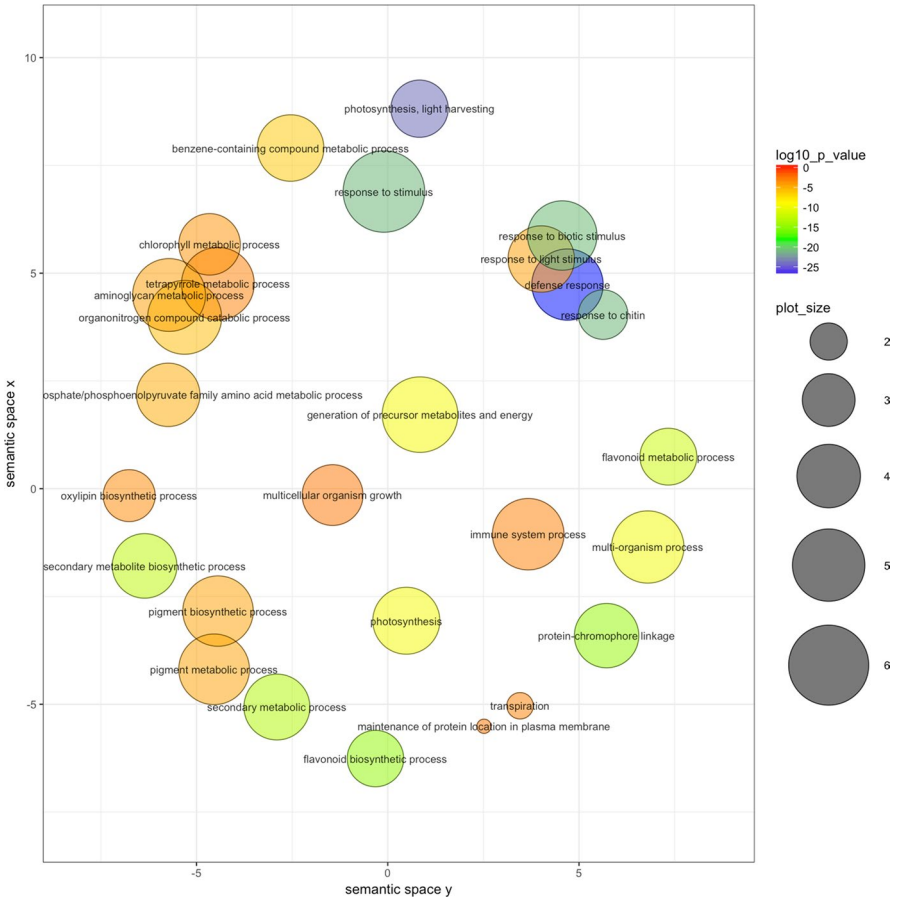

Fig. 2 GO terms of DEGs between the 'Gigantes' and 'Elephantes' using REVIGO software. DEGs were considered those with an absolute log2 fold change > 2 and an adjusted *P*value (P_{adj}) < 0.05 using DESeq2 package

MAPK pathway, and phenylpropanoid biosynthesis pathways (Suppl. Figures 5–8). In particular, the MPK3/6 pathways with its downstream targets, PR1, ERF, ACS6, ChitinaseB, and WRKY33, involved in defense, were all upregulated in 'E' (Suppl. Figure 5). Similarly, from the phenylpropanoid biosynthesis pathway, phenylalanine ammonia lyase, the rate-limiting enzyme of the pathway, as well as many other enzymes involved in the secondary metabolite biosynthesis and conversion, was upregulated in 'E' (Suppl. Figure 6). Similarly, the MPK3/6 pathways with its downstream targets, PR1, ERF, ACS6, ChitinaseB, and WRKY33, involved in defense, were all upregulated in 'E.' Regarding plant hormone signaling, an upregulation was observed for both auxin positive regulators SAUR and GH3 and the negative regulator AUX/IAA, the ABA receptors PYR/PYL, and their downstream kinases SnRK2, as well as components

of the salicylic and jasmonic acid biosynthesis and signaling pathways (Suppl. Figure 7).

Transcription Factor Annotation of DEGs

As a further step, a BLASTX search was conducted to identify transcripts encoding transcription factors (TFs). TFs are key regulators of gene expression associated with various aspects of plant development, stress tolerance, and pathogen resistance. The number of members of each transcription factor family was not drastically different between 'G' and 'E' with a few exceptions. The most populous families were AP2-EREBP, WRKY, AUX/IAA, Homeodomain Binding, and bHLH. Significant differences were observed for the AP2-EREBP, WRKY, and MYB-related which were found in higher counts in 'E,' while bHLH, NAC, GRAS, ARID, and SBP TFs were more numerous in 'G' (Fig. 3). Regarding family members of TFs differentially regulated between the two landraces, members of the AP2-EREBP, ARF, BES1, G2-like, TCP, Trihelix families were significantly upregulated in 'G' (Fig. 4a), and members of AP2-EREBP, WRKY, NAC, bHLH families were upregulated in 'E' (Fig. 4b).

Top DEGs Between 'Gigantes' and 'Elephantes' Landraces

Plotting the DEGs between the two landraces (Fig. 5, Supplementary Table 1) indicated that the genes with the highest expression in 'E' were chalcone isomerase, a pathogenesis-related protein, and the Lon protease, which are key regulatory genes related to secondary metabolism and pathogenesis response. Other highly expressed genes included

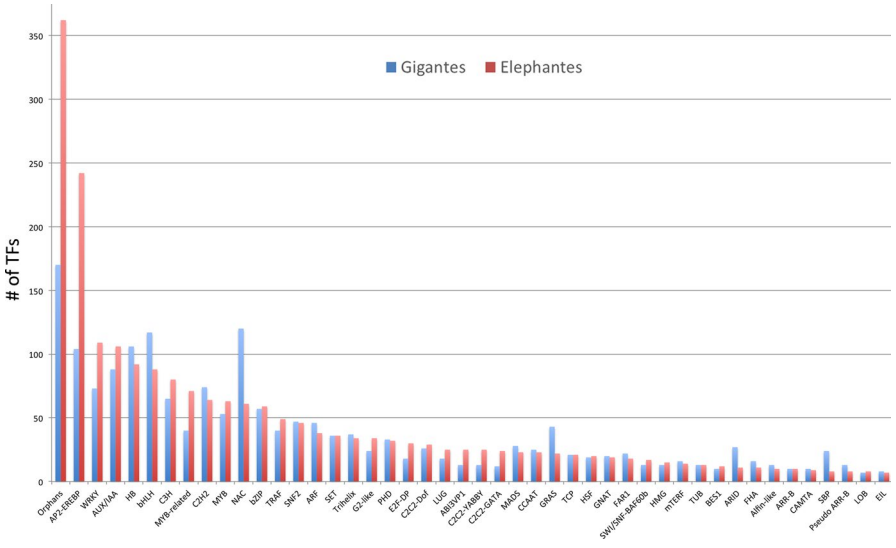

Fig. 3 Distribution of cv. 'Gigantes' and 'Elephantes' transcripts in different transcription factor families

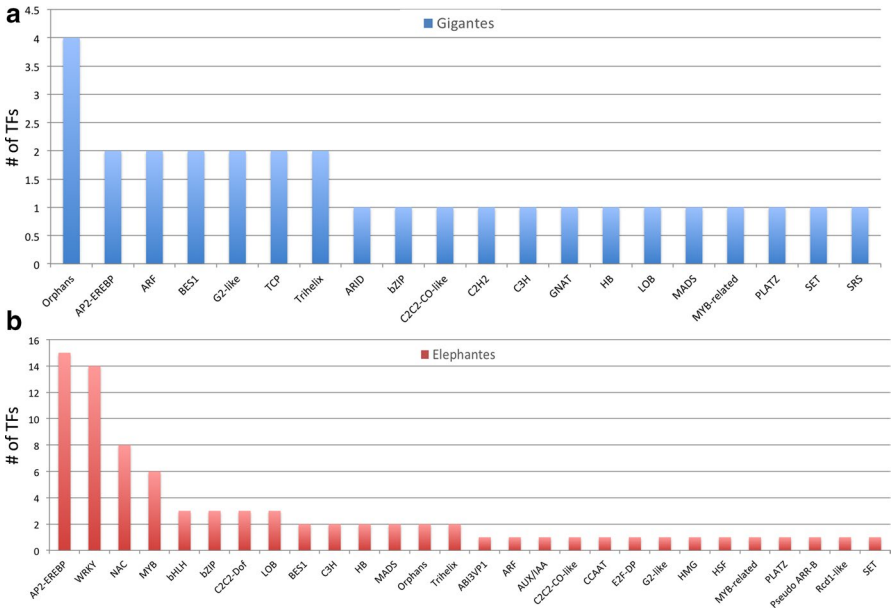

Fig. 4 Transcription factor family members of DEGs upregulated in **a** 'Gigantes' and **b** 'Elephantes'

several pathogenesis, protein turnover, and disease-resistance genes; receptor kinases, ethylene, and jasmonic acid signaling genes; as well as a FRIGIDA-like gene, which regulates flowering; and an α-amylase gene, which catalyzes starch breakdown. On the other hand, DEGs that exhibited higher expression in 'G' compared to 'E' included the *Asr1* gene, involved in abiotic stress tolerance, subtilisin, and patatin-like genes involved in lipid metabolism and cell death execution, chloroplast-located genes which are involved in lipid metabolism as well as albumin-2, a storage protein.

Furthermore, genes that are involved in the regulation of seed size, as supported by functional studies in legumes (Du et al. 2017; Ge et al. 2016; Gu et al. 2017), *Arabidopsis,* and other crops (Li and Li 2015, 2016) were examined. No homologues transcripts were found for the *BIG SEEDS 1* gene, which encodes for a TIFY family transcriptional regulator, a homologue of *ArabidopsisPEAPOD1* gene. The two landraces showed no significant differences in the expression of *WRKY15*, which was correlated to seed size increase in soybean. On the other hand, the larger seed landrace 'E' exhibited 4.2- and 4.1-fold higher expression of the long-chain acyl-CoA synthetase homologue *LACS4* and the histidine kinase homologue *AHK3,* respectively, and -5.4 lower expression of the histidine-containing phosphotransfer protein homologue, *AHP1*, compared to 'G' landrace.

SSRs Identification

The transcriptomes were mined for SSRs (Simple Sequence Repeat) motifs. The range of SSRs length varied between mono- and hexa- nucleotides. The most

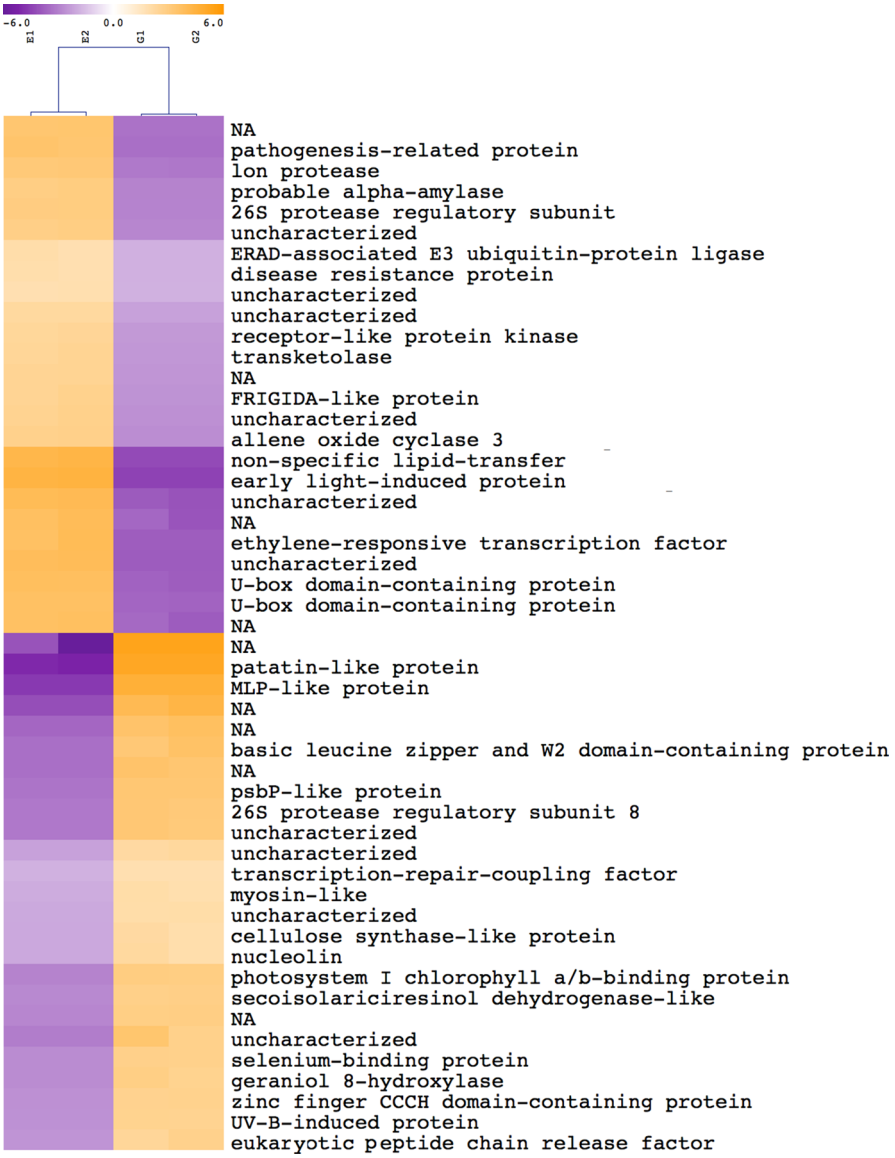

Fig. 5 Heat map plot of the top 50 DEGs between the 'Gigantes' and 'Elephantes'

represented SSRs motifs were trinucleotides for both landraces. The total number of estimated SSR in 'G' was 3956 SSRs in 3391 sequences for a total of 37,379 sequences examined, whereas in 'E' 4322 SSRs in 3588 sequences for a total of 37,774 sequences examined.

Trinucleotide repeats were the most commonly found repetitions in our transcriptome accounting for 41.9% in 'G' and 40.1% in 'E' of the SSRs identified.

The next most abundant class was dinucleotide repeats (29.1% in both landraces) which were closely followed by mononucleotide repeats (26.5% and 28.4% for 'G' and 'E,' respectively). Quad-, penta-, and hexa-nucleotide repeats were observed in very low frequencies (1.3%, 0.4%, and 0.9%, respectively, for 'G,' and 1.2%, 0.5%, and 0.8%, respectively, for 'E,' Fig. 6a).

Among the dinucleotide motifs, the AG motif was the most abundant in both landraces, followed by the TC, GA, and CT repeat motifs (Fig. 6b). Regarding trinucleotide repeats, the most abundant motif in both landraces was GAA with 200 occurrences in 'G' compared to 172 in 'E.' It was followed by the CTT, TTC, TCT, AGA, and AAG motifs which were observed in higher numbers in 'E' compared to 'G' (Fig. 6c).

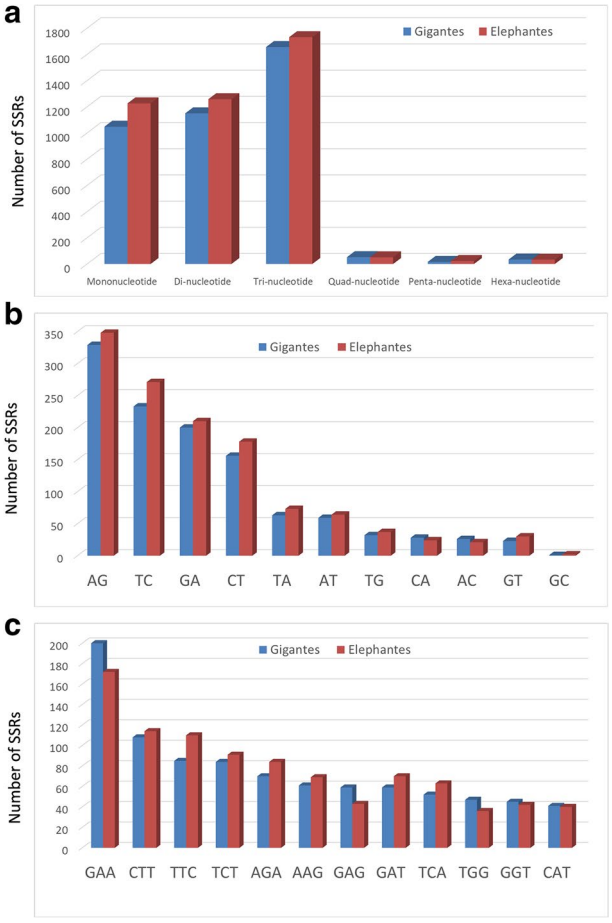

Fig. 6 SSR analysis of transcriptome sequence data of the 'Gigantes' and 'Elephantes.' **a** The number of mono-, di-, tri, quad-, penta-, and hexa-nucleotide SSR motif repeats, **b** the number of each type for dinucleotide and **c** trinucleotide SSRs in 'Gigantes' and 'Elephantes'

Validation of the identified SSR markers was performed using the HRM technique. Among the 10 HRM primers employed, three primer pairs could significantly distinguish the 'G' and 'E' landraces, thus contributing to their discrimination and authentication. The HRM primers employed in this study have been listed in Supplementary Table 2.

SNPs Identification

A total of 19,281 putative SNPs were identified when comparing the transcriptome of 'G' and 'E' landraces (Supplementary Table 2). Among those SNPs, 13,980 were located within the CDS of the assembled transcripts, while 5186 were located in untranslated regions. Of the 13,980 coding SNPs, 7783 were synonymous SNPs, whereas 3161 (22.6%) were non-synonymous SNPs and causing amino acid alterations in protein sequences.

After applying stringent filtering criteria such as non-synonymous, $>100\times$ coverage, 100% allele frequency for each landrace, and 1SNP/transcript, a total of 392 high-quality SNPs were identified. Many of these SNPs result in amino acid alterations with very distinct properties, such as the hydrophilic Asparagine and Histidine being replaced by the hydrophobic Tyrosine, and potential functional consequences.

In addition, to validate the identified high-quality SNPs and their potential use as markers, 19 high-resolution melting (HRM) marker primer sets were designed and the HRM assay was employed in both landraces. Among the 19 HRM primers employed, 16 primers significantly distinguished 'G' and 'E' landraces. The HRM primers employed in this study have been listed in the Supplementary Table 1. The representative HRM melt curves obtained for the parents with the heterozygous SNP variation of G/A and C/A have been illustrated in Fig. 7. Thus, the current SNP-HRM platform may provide a suitable approach for the genotyping purposes in *P. coccineus*.

Discussion

Runner bean is the third most important crop of the *Phaseolus* genus. Since its introduction to Europe it has been cultivated by groups of growers in different geographical pockets. In these regions, different culinary preferences occur, as for example, the preference in southern Europe for white seeds, with higher total sugars, and lower protein and starch content (Santalla et al. 2004), resulting in the generation of cultivars with distinct traits (Rodriguez et al. 2013). This increases the need for detailed genetic characterization, which will both enable runner bean genetic exploitation for yield, resistance, and quality improvement and the establishment of unambiguous identification and labeling of runner bean products. Here is an increased demand for products with specific quality characteristics (Mavromatis et al. 2012).

When comparing the transcriptome annotation between the two *P. coccineus* landraces, no significant differences stood out. This is not surprising, considering the fact that the two landraces have similar phenotypes, differing only in seed size.

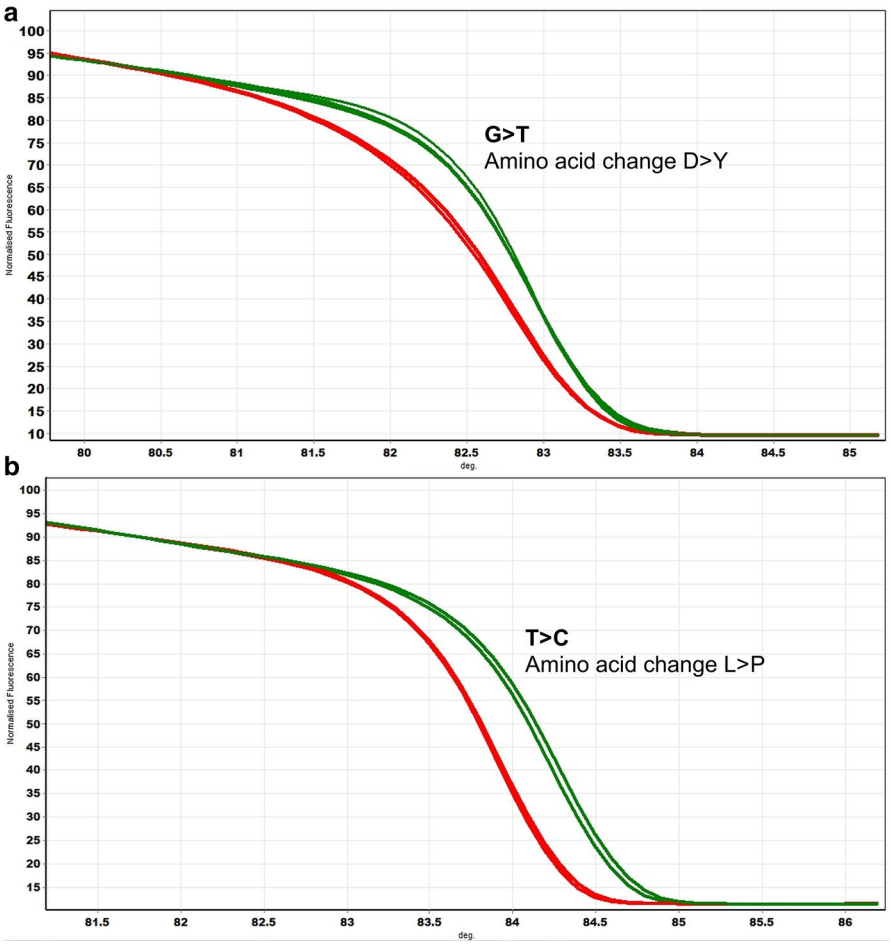

Fig. 7 HRM melt curves obtained from candidate SNPs between 'Gigantes' (red color) and 'Elephantes' (green color) illustrating the G/T and T/C SNP variation (Color figure online)

However, when assessing the annotation of DEGs, a predominant shift towards defense regulatory pathways and secondary metabolism was observed in 'E.' This was explicitly evident considering the phenylpropanoid biosynthetic pathway (Hao et al. 2016), with key regulatory enzymes being upregulated in 'E.' GO categories overrepresented in 'E' were significantly overlapping with those observed in a disease-resistant *Phaseolus vulgaris* cultivar (Padder et al. 2016). This response was also reflected in numerous TFs differentially regulated in 'E' including the WRKY, NAC, MYB, and bHlH families, which are involved in defense responses and secondary metabolism (Wang et al. 2016; Nuruzzaman et al. 2013; Ambawat et al. 2013). On the other hand, genes involved in abiotic stress tolerance and the regulation of cell death genes were upregulated in 'G' compared to 'E.' Specifically the ASR1, which was highly upregulated in 'G,' is involved in abscisic acid signaling

and abiotic stress tolerance and its overexpression in tomato increased survival under dehydration (Golan et al. 2014). *P. coccineus* has been the donor for both abiotic stress-tolerance and pathogen-resistance traits such as cold tolerance (Rodiño et al. 2007), and resistance to foliar and soil-borne fungi (Schwartz et al. 2006; Schwember et al. 2017). Our results indicate that both landraces can be donors either of biotic resistance ('E') or abiotic stress tolerance ('G'), and should be further evaluated and incorporated in breeding programs. No firm conclusions could be drawn by the transcriptome analysis on the differential regulation of seed size between the two landraces. A *BIG SEEDS* 1 gene homologue could not be identified in the transcriptome of both landraces. The loss of function of this gene results in larger seeds in soybean (Ge et al. 2016), thus absence from both landraces transcriptome could be linked to the large size of their seeds, but is not a differentiating factor between them. Significant expression differences in the landraces were observed for members of the Histidine kinase and Histidine-containing phosphotransfer protein gene families. Multiple gene mutants of both families have altered seed size (Hutchison et al. 2006; Li et al. 2013), thus they should be further functionally characterized to confirm their role in seed size regulation.

The application of Next-Generation Sequencing techniques, especially transcriptome assembly with RNA-seq, has allowed the development of large molecular marker datasets even in species which have few genomic resources (Torre et al. 2016; Tanwar et al. 2017; Chen et al. 2016). They enable the identification of thousands of polymorphic loci which allow downstream applications such as molecular authentication (Madesis et al. 2014) and genomics-assisted breeding (Kole et al. 2015). Our work initiates a step towards the expansion of runner bean genomic resources by transcriptome sequencing and de novo assembly of the Greek landraces 'G' and 'E,' with distinct types of seeds. It vastly expands the number of polymorphic markers within *P. coccineus*. We discovered thousands of SSR and SNP polymorphisms that can be used in the construction of genetic maps, to tag important agronomic traits in genomics-assisted breeding and enable the robust authentication of different runner bean cultivars.

Comparing the transcriptomes of the two landraces, 19,281 putative SNPs were identified of which almost a significant portion was non-synonymous. The recent completion of whole-genome sequencing in *Phaseolus vulgaris* has revealed that when compared with other genera within the *Leguminosae* family such as *Glycine*, the genus *Phaseolus* has a higher mutation rate (Schmutz et al. 2014), which could be responsible for the high number and high proportion of non-synonymous SNPs. This high number of non-synonymous SNPs indicates a significant pool of variation in functional traits that can be exploited with genomics-assisted breeding.

Our results demonstrate the power of NGS approaches for polymorphic marker mining. Previous studies undertaken with AFLP and SSR markers indicated that the genetic variation of European accessions of *P.coccineus* was low (Acampora et al. 2007; Boczkowska et al. 2012). This relatively low genetic variation has been related to the limited genetic diversity introduced into Europe, though differences have been observed between different regions in Europe. For instance, accessions from the Iberian Peninsula and Italy exhibited higher variation compared to accessions from central-northern Europe (Spataro et al. 2011; Rodriguez

et al. 2013). The relatively high-confidence SNPs resulting in non-synonymous amino acid changes identified in our study can be utilized in a more detailed screening of European accessions for a comprehensive evaluation of natural diversity, tightly linked to functional traits. They can also enable genome-wide association studies (GWAS) for the identification of genes with underlying important agronomic and adaptive traits of interest.

SSR mining provides another important resource for the identification of multi-allelic genetic polymorphisms which complement SNP markers. Coding region-derived SSRs have a significant level of polymorphism in legumes (Vatanparast et al. 2016). Trinucleotide motifs were predominant in the transcriptome of both 'G' and 'E' landraces, followed by dinucleotide motifs. This pattern has been repeatedly observed in *Phaseolus vulgaris* (Kalavacharla et al. 2011), other legumes (Wang et al. 2014; Tanwar et al. 2017), and other plant species (Torre et al. 2016; Gramazio et al. 2016). The trinucleotide repeats have been shown to be stably inherited (Yang et al. 2012) as they pose a lower risk for deleterious frameshift mutations in translated regions (Metzgar et al. 2000).

The most populous dinucleotide SSR motif AG in the *P. coccineus* transcriptome was identical to that observed in other species from the legumes family such as common bean, winged bean, adzuki bean, and rice bean which show the predominance of AG/CT dinucleotide (Chen et al. 2016; Vatanparast et al. 2016; Kalavacharla et al. 2011), while the most numerous GAA trinucleotide repeat in *P. coccineus* was distinct to the AAG/CTT trinucleotide repeat observed in the species mentioned above. The AG dinucleotide motifs were found to be most abundant in 5′ untranslated regions in *Medicago* (Mun et al. 2006), and can therefore have a regulatory role on transcription by affecting the binding affinity of TFs.

Conclusively, our study significantly expands the number of polymorphic markers available for *P. coccineus,* enabling genetic characterization, authentication, and effective breeding approaches to enhance and to distinguish the quality attributes of runner beans. These transcriptome-derived functional markers, especially SSRs, can be tested in other runner bean cultivars, for the identification of alleles linked with traits of interest, which can then be incorporated or pyramided to cultivars to shape desirable characteristics. Together with the recent discovery of twenty-four domestication-related SNPs, most of them unannotated, as well as SNPs related to cultivar diversification and natural selection (Guerra-García et al. 2017), this information can lead to functional genomics approaches such as gene editing using CRISPR/Cas (Puchta 2017; Shan et al. 2014) to re-shape domestication and to maximize the agronomic potential of runner beans. Finally, runner bean pool of allelic variation can be the driver for *Phaseolus vulgaris* improvement both for yield and resistance traits (Schwember et al. 2017) in order to face the challenges posed by current climatic changes and consumer preferences. Towards this end, identification and mapping of runner bean diversity is of paramount importance.

Acknowledgements This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (Project Code: T1EDK-04718).

References

Acampora A, Ciaffi M, De Pace C, Paolacci AR, Tanzarella OA (2007) Pattern of variation for seed size traits and molecular markers in Italian germplasm of *Phaseolus coccineus* L. Euphytica 157(1):69–82. https://doi.org/10.1007/s10681-007-9397-3

Acosta-Gallegos JA, Kelly JD, Gepts P (2007) Prebreeding in common bean and use of genetic diversity from wild germplasm. Crop Sci 47(Supplement 3):S44–S59

Ambawat S, Sharma P, Yadav NR, Yadav RC (2013) MYB transcription factor genes as regulators for plant responses: an overview. Physiol Mol Biol Plants 19(3):307–321

Bhawna G, Bonthala VS, Gajula MNVP (2016) PvTFDB: a *Phaseolus vulgaris* transcription factors database for expediting functional genomics in legumes. Database 2016:baw114114. https://doi.org/10.1093/database/baw114

Bitocchi E, Rau D, Bellucci E, Rodriguez M, Murgia ML, Gioia T, Santo D, Nanni L, Attene G, Papa R (2017) Beans (*Phaseolus* ssp.) as a model for understanding crop evolution. Front Plant Sci. https://doi.org/10.3389/fpls.2017.00722

Boczkowska M, Bulińska-Radomska Z, Nowosielski J (2012) AFLP analysis of genetic diversity in five accessions of Polish runner bean (*Phaseolus coccineus* L.). Genet Resour Crop Evol 59(4):473–478. https://doi.org/10.1007/s10722-012-9798-6

Chen H, Chen X, Tian J, Yang Y, Liu Z, Hao X, Wang L, Wang S, Liang J, Zhang L (2016). Development of gene-based SSR markers in rice bean (*Vigna umbellata* L.) based on transcriptome data. PLoS ONE 11(3):e0151040

Du J, Wang S, He C, Zhou B, Ruan Y-L, Shou H (2017) Identification of regulatory networks and hub genes controlling soybean seed set and size using RNA sequencing analysis. J Exp Bot 68(8):1955–1972

Escalante AM, Coello G, Eguiarte LE, Pinero D (1994) Genetic structure and mating systems in wild and cultivated populations of *Phaseolus coccineus* and *P. vulgaris* (Fabaceae). Am J Bot 81:1096–1103

Ganopoulos I, Bosmali I, Madesis P, Tsaftaris A (2012) Microsatellite genotyping with HRM (High Resolution Melting) analysis for identification of the PGI common bean variety Plake Megalosperma Prespon. Eur Food Res Technol 234(3):501–508

Ganopoulos I, Tsaballa A, Xanthopoulou A, Madesis P, Tsaftaris A (2013) Sweet cherry cultivar identification by high-resolution-melting (HRM) analysis using gene-based SNP markers. Plant Mol Biol Rep 31(3):763–768

Ge L, Yu J, Wang H, Luth D, Bai G, Wang K, Chen R (2016) Increasing seed size and quality by manipulating *BIG SEEDS1* in legume species. Proc Natl Acad Sci USA 113(44):12414–12419

Golan I, Dominguez PG, Konrad Z, Shkolnik-Inbar D, Carrari F, Bar-Zvi D (2014) Tomato *ABSCISIC ACID STRESS RIPENING* (*ASR*) gene family revisited. PLoS ONE 9(10):e107117. https://doi.org/10.1371/journal.pone.0107117

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29(7):644–652

Gramazio P, Blanca J, Ziarsolo P, Herraiz F, Plazas M, Prohens J, Vilanova S (2016). Transcriptome analysis and molecular marker discovery in *Solanum incanum* and *S. aethiopicum*, two close relatives of the common eggplant (*Solanum melongena*) with interest for breeding. BMC Genomics 17(1):300

Gu Y, Li W, Jiang H, Wang Y, Gao H, Liu M, Chen Q, Lai Y, He C (2017) Differential expression of a *WRKY* gene between wild and cultivated soybeans correlates to seed size. J Exp Bot 68(11):2717–2729

Guerra-García A, Suárez-Atilano M, Mastretta-Yanes A, Delgado-Salinas A, Piñero D (2017) Domestication genomics of the open-pollinated scarlet runner bean (Phaseolus coccineus L.). Front Plant Sci 8:1891. https://doi.org/10.3389/fpls.2017.01891

Hao C, Xia Z, Fan R, Tan L, Hu L, Wu B, Wu H (2016) De novo transcriptome sequencing of black pepper (*Piper nigrum* L.) and an analysis of genes involved in phenylpropanoid metabolism in response to *Phytophthora capsici*. BMC Genomics 17(1):822

Hutchison CE, Li J, Argueso C, Gonzalez M, Lee E, Lewis MW, Maxwell BB, Perdue TD, Schaller GE, Alonso JM, Ecker JR, Kieber JJ (2006) The *Arabidopsis* histidine phosphotransfer proteins are redundant positive regulators of cytokinin signaling. Plant Cell 18(11):3073–3087. https://doi.org/10.1105/tpc.106.045674

Hyten DL, Song Q, Fickus EW, Quigley CV, Lim J-S, Choi I-Y, Hwang E-Y, Pastor-Corrales M, Cregan PB (2010) High-throughput SNP discovery and assay development in common bean. BMC Genomics 11(1):475. https://doi.org/10.1186/1471-2164-11-475

Kage U, Kumar A, Dhokane D, Karre S, Kushalappa AC (2016) Functional molecular markers for crop improvement. Crit Rev Biotechnol 36(5):917–930

Kalavacharla V, Liu Z, Meyers BC, Thimmapuram J, Melmaiee K (2011) Identification and analysis of common bean (*Phaseolus vulgaris* L.) transcriptomes by massively parallel pyrosequencing. BMC Plant Biol 11(1):135

Kole C, Muthamilarasan M, Henry R, Edwards D, Sharma R, Abberton M, Batley J, Bentley A, Blakeney M, Bryant J (2015) Application of genomics-assisted breeding for generation of climate resilient crops: progress and prospects. Front Plant Sci 6:563

Kumar S, Banks TW, Cloutier S (2012) SNP discovery through next-generation sequencing and its applications. Int J Plant Genomics 2012:831460

Li N, Li Y (2015) Maternal control of seed size in plants. J Exp Bot 66(4):1087–1097

Li N, Li Y (2016) Signaling pathways of seed size control in plants. Curr Opin Plant Biol 33:23–32

Li J, Nie X, Tan JLH, Berger F (2013) Integration of epigenetic and genetic controls of seed size by cytokinin in *Arabidopsis*. Proc Natl Acad Sci USA 110(38):15479–15484. https://doi.org/10.1073/pnas.1305175110

Li J, Dai X, Zhuang Z, Zhao PX (2016) LegumeIP 2.0—a platform for the study of gene function and genome evolution in legumes. Nucleic Acids Res 44(D1):D1189–D1194. https://doi.org/10.1093/nar/gkv1237

Madesis P, Ganopoulos I, Sakaridis I, Argiriou A, Tsaftaris A (2014) Advances of DNA-based methods for tracing the botanical origin of food products. Food Res Int 60:163–172

Mavromatis A, Arvanitoyannis I, Chatzitheodorou V, Kaltsa A, Patsiaoura I, Nakas C (2012) A comparative study among landraces of *Phaseolus vulgaris* L. and *P. coccineus* L. based on molecular, physicochemical and sensory analysis for authenticity purposes. Sci Hortic 144:10–18

Metzgar D, Bytof J, Wills C (2000) Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Res 10(1):72–80

Mun J-H, Kim D-J, Choi H-K, Gish J, Debellé F, Mudge J, Denny R, Endré G, Saurat O, Dudez A-M (2006) Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. Genetics 172(4):2541–2555

Nuruzzaman M, Sharoni AM, Kikuchi S (2013) Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Front Microbiol 4:248

Padder BA, Kamfwa K, Awale HE, Kelly JD (2016) Transcriptome profiling of the *Phaseolus vulgaris–Colletotrichum lindemuthianum* pathosystem. PLoS ONE 11(11):e0165823. https://doi.org/10.1371/journal.pone.0165823

Puchta H (2017) Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. Curr Opin Plant Biol 36:1–8

Rendón-Anaya M, Montero-Vargas JM, Saburido-Álvarez S, Vlasova A, Capella-Gutierrez S, Ordaz-Ortiz JJ, Aguilar OM, Vianello-Brondani RP, Santalla M, Delaye L, Gabaldón T, Gepts P, Winkler R, Guigó R, Delgado-Salinas A, Herrera-Estrella A (2017) Genomic history of the origin and domestication of common bean unveils its closest sister species. Genome Biol 18(1):60. https://doi.org/10.1186/s13059-017-1190-6

Rodiño AP, Lema M, Pérez-Barbeito M, Santalla M, De Ron AM (2007) Assessment of runner bean (*Phaseolus coccineus* L.) germplasm for tolerance to low temperature during early seedling growth. Euphytica 155(1–2):63–70

Rodriguez M, Rau D, Angioi SA, Bellucci E, Bitocchi E, Nanni L, Knüpffer H, Negri V, Papa R, Attene G (2013) European *Phaseolus coccineus* L. landraces: population structure and adaptation, as revealed by cpSSRs and phenotypic analyses. PLoS ONE 8(2):e57337. https://doi.org/10.1371/journal.pone.0057337

Rubiales D, Mikic A (2015) Introduction: legumes in sustainable agriculture. Taylor & Francis, Boca Raton

Santalla M, Monteagudo A, González A, De Ron A (2004) Agronomical and quality traits of runner bean germplasm and implications for breeding. Euphytica 135(2):205–215

Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu S, Song Q, Chavarro C, Torres-Torres M, Geffroy V, Moghaddam SM, Gao D, Abernathy B, Barry K, Blair M, Brick MA, Chovatia M, Gepts P, Goodstein DM, Gonzales M, Hellsten U, Hyten DL, Jia G, Kelly JD, Kudrna D, Lee R, Richard MMS, Miklas PN, Osorno JM, Rodrigues J, Thareau V, Urrea

CA, Wang M, Yu Y, Zhang M, Wing RA, Cregan PB, Rokhsar DS, Jackson SA (2014) A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet 46(7):707–713. https://doi.org/10.1038/ng.3008

Schröder S, Mamidi S, Lee R, McKain MR, McClean PE, Osorno JM (2016) Optimization of genotyping by sequencing (GBS) data in common bean (*Phaseolus vulgaris* L.). Mol Breed 36(1):6. https://doi.org/10.1007/s11032-015-0431-1

Schwartz HF, Otto K, Terán H, Lema M, Singh SP (2006) Inheritance of white mold resistance in *Phaseolus vulgaris* × *P. coccineus* crosses. Plant Dis 90(9):1167–1170

Schwember AR, Carrasco B, Gepts P (2017) Unraveling agronomic and genetic aspects of runner bean (*Phaseolus coccineus* L.). Field Crops Res 206:86–94. https://doi.org/10.1016/j.fcr.2017.02.020

Shan Q, Wang Y, Li J, Gao C (2014) Genome editing in rice and wheat using the CRISPR/Cas system. Nature Protoc 9(10):2395–2410

Sicard D, Nanni L, Porfiri O, Bulfon D, Papa R (2005) Genetic diversity of *Phaseolus vulgaris* L. and *P. coccineus* L. landraces in central Italy. Plant Breed 124(5):464–472. https://doi.org/10.1111/j.1439-0523.2005.01137.x

Spataro G, Tiranti B, Arcaleni P, Bellucci E, Attene G, Papa R, Zeuli PS, Negri V (2011) Genetic diversity and structure of a worldwide collection of *Phaseolus coccineus* L. Theor Appl Genet 122(7):1281–1291

Tanwar UK, Pruthi V, Randhawa GS (2017) RNA-seq of guar (*Cyamopsis tetragonoloba*, L. Taub.) leaves: de novo transcriptome assembly, functional annotation and development of genomic resources. Front Plant Sci 8:91. https://doi.org/10.3389/fpls.2017.00091

Torre S, Tattini M, Brunetti C, Guidi L, Gori A, Marzano C, Landi M, Sebastiani F (2016) De novo assembly and comparative transcriptome analyses of red and green morphs of sweet basil grown in full sunlight. PLoS ONE 11(8):e0160370

Vatanparast M, Shetty P, Chopra R, Doyle JJ, Sathyanarayana N, Egan AN (2016) Transcriptome sequencing and marker development in winged bean (*Psophocarpus tetragonolobus*; *Leguminosae*). Sci Rep 6:29070

Vaz Patto MC, Amarowicz R, Aryee AN, Boye JI, Chung H-J, Martín-Cabrejas MA, Domoney C (2015) Achievements and challenges in improving the nutritional quality of food legumes. Crit Rev Plant Sci 34(1–3):105–143

Vlasova A, Capella-Gutiérrez S, Rendón-Anaya M, Hernández-Oñate M, Minoche AE, Erb I, Câmara F, Prieto-Barja P, Corvelo A, Sanseverino W, Westergaard G, Dohm JC, Pappas GJ, Saburido-Alvarez S, Kedra D, Gonzalez I, Cozzuto L, Gómez-Garrido J, Aguilar-Morón MA, Andreu N, Aguilar OM, Garcia-Mas J, Zehnsdorf M, Vázquez MP, Delgado-Salinas A, Delaye L, Lowy E, Mentaberry A, Vianello-Brondani RP, García JL, Alioto T, Sánchez F, Himmelbauer H, Santalla M, Notredame C, Gabaldón T, Herrera-Estrella A, Guigó R (2016) Genome and transcriptome analysis of the Mesoamerican common bean and the role of gene duplications in establishing tissue and temporal specialization of genes. Genome Biol 17(1):32. https://doi.org/10.1186/s13059-016-0883-6

Wang Z, Yu G, Shi B, Wang X, Qiang H, Gao H (2014) Development and characterization of simple sequence repeat (SSR) markers based on RNA-sequencing of *Medicago sativa* and in silico mapping onto the *M. truncatula* genome. PLoS ONE 9(3):e92029

Wang N, Xia E-H, Gao L-Z (2016) Genome-wide analysis of WRKY family of transcription factors in common bean, *Phaseolus vulgaris*: chromosomal localization, structure, evolution and expression divergence. Plant Gene 5:22–30. https://doi.org/10.1016/j.plgene.2015.11.003

Xanthopoulou A, Ganopoulos I, Tsaballa A, Nianiou-Obeidat I, Kalivas A, Tsaftaris A, Madesis P (2014) Summer squash identification by High-Resolution-Melting (HRM) analysis using gene-based EST–SSR molecular markers. Plant Mol Biol Rep 32(2):395–405

Xanthopoulou A, Ganopoulos I, Psomopoulos F, Manioudaki M, Moysiadis T, Kapazoglou A, Osathanunkul M, Michailidou S, Kalivas A, Tsaftaris A (2017) De novo comparative transcriptome analysis of genes involved in fruit morphology of pumpkin cultivars with extreme size difference and development of EST-SSR markers. Gene 622:50–66

Yang T, Bao S, Ford R, Jia T, Guan J, He Y, Sun X, Jiang J, Hao J, Zhang X (2012) High-throughput novel microsatellite marker of faba bean via next generation sequencing. BMC Genomics 13(1):602

Affiliations

Aliki Xanthopoulou[1] · Christos Kissoudis[2] · Ifigeneia Mellidou[3] · Maria Manioudaki[4] · Irene Bosmali[1] · Vasilis Tsakonitis[3] · Apostolos Kalivas[3] · Maslin Osathanunkul[5,6] · Athanasios Tsaftaris[2] · Ioannis Ganopoulos[3] · Panagiotis Madesis[1]

✉ Ioannis Ganopoulos
giannis.ganopoulos@gmail.com

✉ Panagiotis Madesis
pmadesis@certh.gr

[1] Institute of Applied Biosciences, CERTH, Thermi, 570 01 Thessaloniki, Greece

[2] Perrotis College, American Farm School, 570 01 Thessaloniki, Greece

[3] Institute of Plant Breeding and Genetic Resources, ELGO-DEMETER, Thermi, 570 01 Thessaloniki, Greece

[4] Centre for Clinical, Experimental Surgery & Translational Research, Biomedical Research Foundation Academy of Athens, Athens, Greece

[5] Department of Biology, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand

[6] Center of Excellence in Bioresources for Agriculture, Industry and Medicine, Chiang Mai University, Chiang Mai, Thailand