

RESEARCH ARTICLE

Doubled Haploid 'CUDH2107' as a Reference for Bulb Onion (*Allium cepa* L.) Research: Development of a Transcriptome Catalogue and Identification of Transcripts Associated with Male Fertility

Jiffinvir S. Khosa¹, Robyn Lee¹, Sophia Bräuning^{1,2}, Janice Lord², Meeghan Pither-Joyce³, John McCallum^{1,3}, Richard C. Macknight^{1,3*}

1 Department of Biochemistry, University of Otago, Dunedin, New Zealand, **2** Department of Botany, University of Otago, Dunedin, New Zealand, **3** New Zealand Institute for Plant & Food Research, Lincoln, New Zealand

* richard.macknight@otago.ac.nz



OPEN ACCESS

Citation: Khosa JS, Lee R, Bräuning S, Lord J, Pither-Joyce M, McCallum J, et al. (2016) Doubled Haploid 'CUDH2107' as a Reference for Bulb Onion (*Allium cepa* L.) Research: Development of a Transcriptome Catalogue and Identification of Transcripts Associated with Male Fertility. PLoS ONE 11(11): e0166568. doi:10.1371/journal.pone.0166568

Editor: Xiang Jia Min, Youngstown State University, UNITED STATES

Received: July 12, 2016

Accepted: October 31, 2016

Published: November 18, 2016

Copyright: © 2016 Khosa et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All the sequence data is available at NCBI, <http://www.ncbi.nlm.nih.gov/bioproject/>, Biosample numbers 5898004-9.

Funding: Jiffinvir S. Khosa's PhD research is supported by an Indian Council of Agricultural Research (ICAR) International fellowship, an Otago University doctoral scholarship, and a New Zealand Onion Industry Postgraduate Scholarship.

Abstract

Researchers working on model plants have derived great benefit from developing genomic and genetic resources using 'reference' genotypes. Onion has a large and highly heterozygous genome making the sharing of germplasm and analysis of sequencing data complicated. To simplify the discovery and analysis of genes underlying important onion traits, we are promoting the use of the homozygous double haploid line 'CUDH2107' by the onion research community. In the present investigation, we performed transcriptome sequencing on vegetative and reproductive tissues of CUDH2107 to develop a multi-organ reference transcriptome catalogue. A total of 396 million 100 base pair paired reads was assembled using the Trinity pipeline, resulting in 271,665 transcript contigs. This dataset was analysed for gene ontology and transcripts were classified on the basis of putative biological processes, molecular function and cellular localization. Significant differences were observed in transcript expression profiles between different tissues. To demonstrate the utility of our CUDH2107 transcriptome catalogue for understanding the genetic and molecular basis of various traits, we identified orthologues of rice genes involved in male fertility and flower development. These genes provide an excellent starting point for studying the molecular regulation, and the engineering of reproductive traits.

Introduction

Bulb onion (*Allium cepa* L.) is a monocot vegetable crop grown for edible bulbs and has economic importance worldwide. The onion research community would benefit from improved onion genomic resources [1, 2]. In recent years, next generation sequencing technologies have been used in crop plants to generate genomic and transcriptomic data sets in a cost and time-

Competing Interests: The authors have declared that no competing interests exist.

effective manner. The use of RNA sequencing (RNA-seq) enables researchers to discover genes and molecular markers associated with important traits for their breeding programmes [3, 4]. RNA-seq is a particularly attractive approach in non-model crops that have large genomes, where genomic sequencing is complex and expensive. To aid downstream analysis and avoid detection of false SNPs, a good quality transcriptome assembly is essential [3, 5, 6]. However in species that are highly heterozygous, such as bulb onion, the development of a high quality reference transcriptome is challenging, as it is hard to distinguish the transcripts belonging to different members of a gene family from allelic variants of a particular gene [2, 3, 5]. In out-crossing crop plants, such as bulb onion, the complications of heterozygosity can be overcome by using homozygous double haploids [1, 6]. Double haploid (DH) lines have been developed in bulb onion and have proved useful for various genetic and genomic studies [7–10]. There are many advantages in the use of a common reference double haploid line for genetic and genomic studies by researchers throughout the world. Unfortunately the majority of onion DH lines are neither vigorous nor have good seed production, which complicates wider distribution and usage. In contrast, a set of DH lines developed from a synthetic background at Cornell University [7] have proved to be more widely usable for breeding [11]. We have suggested that ‘CUDH2107’ be employed as a common reference line, as it is relatively vigorous, produces adequate amounts of seed and produces a bulb that stores well [12].

The development of F₁ hybrid onions has transformed the quality and yield of onion production. However, there are concerns that the introduction of F₁ hybrids has reduced the diversity of germplasm being grown. As only two sources of male sterility (CMS-S and CMS-T) have been utilized, it would be desirable to identify additional sources of male sterility in bulb onion [2]. In other plant species, wide hybridization and induced mutagenesis have been utilized to develop male sterile phenotypes [13]. The male sterile mutants often have either abnormal development of sporophytic anther tissues (primarily tapta and meiotic cells) causing lack of pollen or pollen abortion, or have abnormal development of gametophytic anther tissues affecting microspore or pollen grain formation. There is a large body of research into the genetics and molecular mechanisms of male sterility and fertility restoration in other plants, especially monocots, such as rice and maize that could potentially be applied to onion [13–15]. Recently, the CMS-S onion mitochondrial genome was sequenced, leading to the finding that orf725 might be the most plausible candidate gene responsible for inducing male sterility [16]. Further, a gene encoding PMS1, involved in the DNA mismatch repair pathway, was identified as the possible candidate gene regulating fertility restoration [17]. However, the molecular mechanisms of male sterility and fertility restoration in bulb onion is still poorly understood [2].

In this paper, we develop a transcriptome catalogue for ‘CUDH2107’ as a resource for the *Allium* research community [12]. To demonstrate the utility of this data, we identified orthologues of rice genes involved in male fertility and restoration of CMS, which could be useful for studying these processes in bulb onion. This information provides potential targets for the development of novel sources of male sterility for hybrid seed production, by using new genome editing such as CRISPR/cas9 to induce specific mutations in these genes.

Material and Methods

Plant material and transcriptome sequencing

Seed lots of the long day DH bulb onion ‘CUDH2107’ (line CUDH066607 in [11]) were provided by Cornell University (US). Total RNA was extracted from tissue samples pooled from multiple plants grown in tunnel houses at Lincoln New Zealand (latitude 42° S) or in controlled environments. The stages sampled were as follows: leaves (from plants grown in long

day of 16 h light: 8 h dark), floral buds from unexpanded umbels, unopened florets from expanded umbels, open florets with pollen, older flowers and roots. RNA was isolated using a Qiagen RNA extraction kit following the manufacturer's guidelines. Libraries were made using the TruSeq v2 kit (Illumina), and were sequenced on the Illumina HiSeq 2000 platform by NZGL Ltd.

De novo assembly of transcripts

The program fastq_quality_trimmer (FASTX_toolkit, version 0.0.13) was used to trim bases with a quality score less than 30, subsequently any reads containing shorter than 20 bases were removed. The cleaned reads from the CUDH2107 onion tissue libraries were assembled together in a single reference *de novo* assembly using Trinity [18] following the protocol and default parameters [19]. The combined *de novo* assembly is referred to as the 'extensive transcriptome dataset', which was filtered based on a minimum fragments per kb of target transcript length per million (FPKM) value of 0.5 (41 reads per kb). As a result we compiled the 'abundant transcriptome dataset' of highly expressed transcripts, which was used in all the analyses described in this paper. All the sequence data is deposited at NCBI as sequence read archive (S1 Table).

The completeness of the extensive and abundant transcriptomes was assessed based on assembly statistics achieved by running the script 'TrinityStats.pl' [18]. In addition, the eukaryotic Benchmarking Universal Single-copy Orthologs (BUSCOs) dataset (<http://busco.ezlab.org/>, accessed on 20 May 2015) was compared with our abundant transcriptome dataset using BUSCO_v1.1 [20].

Sequence conservation and functional annotation

Standalone BLAST (ncbi-blast-2.2.27+, [21]) was used to perform sequence similarity searches of the current onion transcriptome assembly to a variety of transcriptome assemblies and rice proteins. BLASTN with an E-value cut off of 10^{-4} was used to estimate the sequence conservation among rice and other transcriptomic assemblies of onion, bunching onion, and garlic [8, 10, 22–25]. BLASTX search with an E-value cut off of 10^{-4} was used to compare the peptides encoded by the onion transcripts to rice proteins [26] and the results were used to obtain Gene Ontology (GO) terms for the onion transcripts. This was achieved using GO annotations identifiers from the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/downloads_gad.shtml).

To further identify transcripts potentially coding for full-length peptides, the abundant transcriptome dataset was screened for Open Reading Frames (ORFs) using the ORF-predictor server [27] (<http://proteomics.yzu.edu/tools/OrfPredictor.html>). The resulting predicted peptides were filtered, using custom python and R scripts (available on request), to only retain transcripts with predicted peptides that are at least 100 amino acids long.

Abundance estimation and differential expression analysis

The trinity protocol was followed for abundance estimation, differential expression and hierarchical clustering [19]. Transcript abundance was calculated by first aligning the trimmed reads from each sample to the extensive transcriptome dataset using Bowtie then RSEM [28] was used to estimate abundance of each transcript. The differential transcript expression between different samples was calculated using the Bioconductor package EdgeR [29]. To compare transcriptional profiles across samples, transcripts differentially expressed in at least one pairwise comparison were used to perform hierarchical clustering of transcripts and samples. For the hierarchical clustering, the FPKM values (obtained from RSEM) were \log_2 -transformed

and median-centered. To compare correlation between each sample pair, TMM (Trimmed Mean of M-values) normalized FPKM values were used to obtain a Spearman correlation matrix, then the correlation matrix was hierarchically clustered and visualized as a heat map.

Identification of male fertility genes

The coding DNA sequence of rice flowering genes [26] was used as query to perform BLASTn against bulb onion transcriptome data with an E value cut off of $1e^{-4}$. Top blast hits from bulb onion were translated and use as query sequence in reciprocal BLASTp searches against rice database. The bulb onion contigs retrieving rice genes after reciprocal blast were selected for further analysis. The multiple sequence alignment using amino acid sequences was carried out with GENEIOUS 6.1. The aligned sequences were used for generating trees based on Neighbour Joining Method in the GENEIOUS 6.1 software package. The relative expression of bulb onion genes in different samples was calculated based on FPKM values.

Quantitative real-time PCR

The differential expression of genes involved in flower development was validated using qPCR. Total RNA was isolated from different development stages using Plant RNA Purification Reagent (Invitrogen, USA) following manufacturer's guidelines. Reverse transcription was carried out with 1µg of total RNA using Invitrogen Super Script III following manufacturer's guidelines. Quantitative real time PCR was carried out using 10µL SYBR reaction mixture (Kapa Biosystems) in a Roche Light Cycler 480. Relative gene expression levels were calculated using the $2^{-\Delta\Delta C_T}$ method in Roche LC480 software. Actin and β -tubulin were used as the reference genes. The list of primer sequences used in present investigation was given in [S2 Table](#).

Results and Discussion

Transcriptome of bulb onion and its comparison with other alliums

Illumina sequencing was carried out on cDNA libraries developed from leaves, immature flower heads, unopened flowers, opened flowers with pollen, older flowers, and roots. This resulted in approximately 396 million 100 base pair paired reads. Using Trinity software, cleaned reads were *de novo* assembly into 362,106 contigs representing what we called the 'extensive transcriptome dataset' of bulb onion. Transcripts had a total length of 218.6 Mbp with an average length of 603bp and N50 length of 901bp. We filtered out the low abundance contigs and mostly short contigs from the extensive transcriptome dataset using a minimum FPKM value of 0.5, which represents an average base coverage of 8.2. This resulted in 271,665 highly expressed contigs with an average length of 653bp and N50 length of 1055bp (combined transcript length 177Mbp) ([Fig 1](#); [Table 1](#)). In the abundant transcriptome dataset, 266,427 transcripts were predicted to encode peptides, with 50,220 transcripts encoding peptides that are at least 100 amino acids long.

Only 15% to 49% of the highly abundant transcripts identified by this study were similar to those previously identified in bulb onion [8–10, 23], highlighting the value of our multiple organ transcriptome assembly. The half of the transcripts from this transcriptome assembly were present in an assembly from six week old seedling short and long day onions [23]. Also in garlic, 78% of transcripts generated in multiple organ transcriptome were present in the transcriptome developed from single tissue [22, 24]. Transcriptome assemblies have been developed from other economically important *Allium* species (bunching onion and garlic) and their comparison with bulb onion gives us an idea about the degree of transcriptome

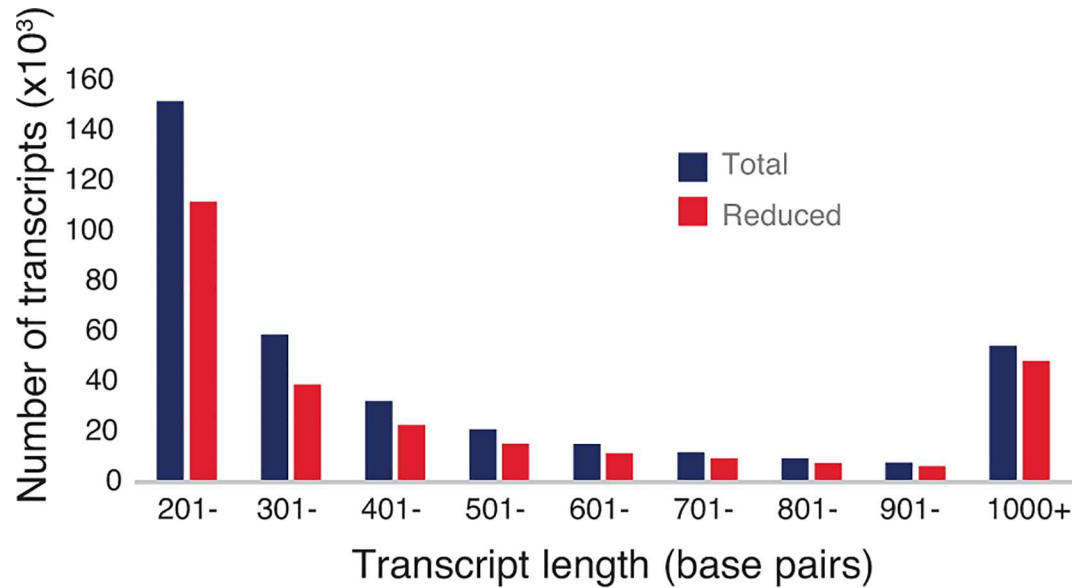


Fig 1. Length distribution of assembled transcripts in the extensive (Total) and abundant (Reduced) transcriptome datasets.

doi:10.1371/journal.pone.0166568.g001

conservation in the genus *Allium* [22, 24, 25]. Only 22% and 10% of bulb onion transcripts were highly similar to transcripts from bunching onion and garlic transcriptomes, respectively. However, those that were similar shared 93% identity with bunching onion and 90% with garlic (Table 2), supporting the fact that bulb onion is more closely related to bunching onion than to garlic [30]. The common transcripts identified between different alliums in the present investigation will be useful for better understanding of *Allium* comparative genomics.

Gene prediction and functional annotation

A total of 56,805 (20.91%) transcripts showed significant hits with rice proteins and shared 56% average identity. To further validate the gene predictions, we used the predicted bulb onion peptides to search KOGs, the core genes from the Benchmarking Universal Single-Copy Orthologues (BUSCOs) pipeline [20]. This search revealed that the transcriptome contains 82% complete BUSCOs (250) and 4% fragmented BUSCOs and indicates a near complete transcriptome.

All predicted bulb onion peptides were functionally annotated following a consensus approach using GO slims from the rice genome annotation database. The bulb onion

Table 1. Statistics of De novo assembly and abundance estimation.

	Extensive (Total)	Abundant (Reduced)
Number of Transcripts	362,106	271,665
Total transcripts length (bp)	218,643,405	177,424,188
Average Transcript length (bp)	603.81	653.10
Median Transcript length (bp)	341	355
Minimum Transcript length (bp)	201	201
Maximum Transcript Length (bp)	20,231	20,231
N50 (bp)	901	1055
GC %	36.96	37.12

doi:10.1371/journal.pone.0166568.t001

Table 2. Sequence conservation between the bulb onion abundant transcriptome dataset and other *Allium* species.

Species	Reference database ^a	Number of Reference sequences	Number and % of hits in reference database	Average Identity (%) of hits	Reference
Bunching onion	274623	54,903	59,923 (22.06)	93.63	25
Bulb onion	246669	367,683	133,547 (49.16)	98.04	23
Bulb onion	238142	128,598	114,151 (42.02)	98.79	8
Bulb onion	175446	33,162	52,297 (19.25)	97.59	10
Bulb onion	175449	26,995	42,264 (15.56)	97.58	10
Garlic	158177	79,143	27,383 (10.08)	90.92	22

^a <http://www.ncbi.nlm.nih.gov/bioproject/>

doi:10.1371/journal.pone.0166568.t002

transcripts were grouped into 95 functional groups (Cellular component, Biological Process and Molecular Function) A similar number of GO terms were found in bulb onion and rice but the number of genes with GO terms in various categories differed (Fig 2A), which might reflect differences in life cycle, development stages and physiological pathways [31]. We found 24 categories within ‘Cellular Components’, 26 within ‘Biological Process’ and 45 within ‘Molecular Function’ categories (Fig 2). The top GO terms for ‘Cellular Component’ were cell (7267), followed by cell wall (5912) and cellular components (4987) (Fig 2B). For ‘Biological Processes’, top GO terms were Abscission (11,710), followed by Anatomical Structure Morphogenesis (11,054) and Behaviour (10,775) (Fig 2C). In the case of ‘Molecular Function’, Binding Domains (6811) was the most abundant GO term, followed by Carbohydrate Binding (5294) and Catalytic Activity (5152) (Fig 2D).

GC content

GC content is a striking characteristic of genome organization and life history of plant species [32]. Bulb onion has a lower GC content than grasses, which might be due to the large genome size found in bulbous geophytes [32, 33]. The average GC content in the present transcriptome dataset is ~38%, whereas bunching onion transcriptome have 40% GC content [25]. The GC content in present investigation is lower than that previously reported based on small EST dataset [33]. This difference might be due to variation in gene length, structure, expression and methylation in these datasets, as these factors affect GC content [34]. Overall our findings confirm the occurrence of low GC content in genus *Allium*.

Tissue specific expression

To study the expression pattern of transcripts across different bulb onion tissues, pairwise comparisons were used to identify transcripts that are differentially expressed in at least one tissue. Using a significance threshold of 0.001 False Discovery Rate and 4-fold change in expression, we determined that there were 17 thousand transcripts differentially expressed among different tissues. ‘Unopened flowers’ and ‘open flowers with pollen’ shared a more similar pattern of expression, with the next most similar sample being ‘older flowers’ (Fig 3). However these samples demonstrated a quite different pattern of expression to that of ‘immature flower heads’ (Fig 3). Transcripts from leaves and roots also showed distinct expression patterns, as they grouped on separate nodes (Fig 3).

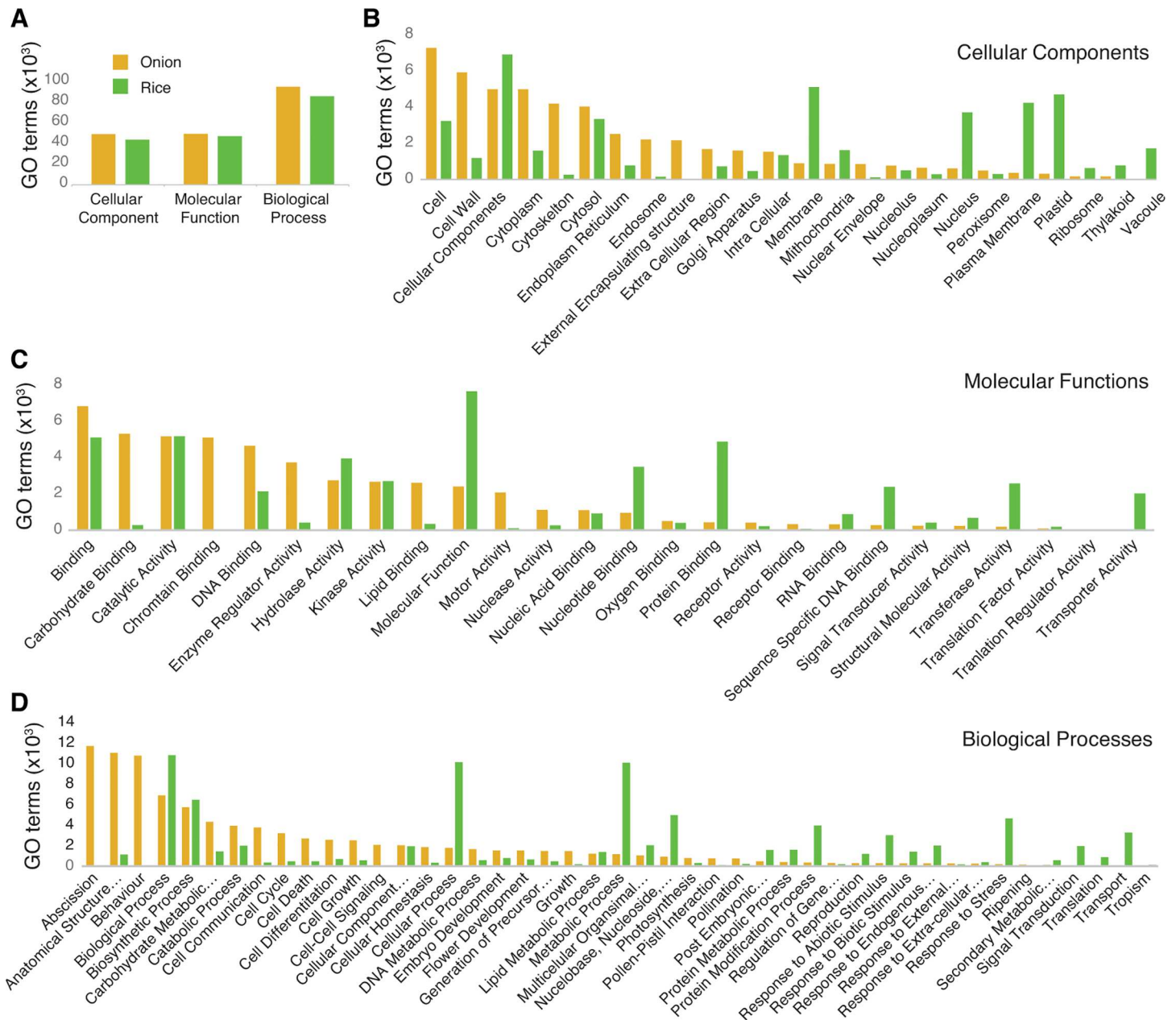


Fig 2. GO terms in bulb onion compared with rice. (A) Total number of GO terms associated with cellular component, molecular function and biological process in onion (brown) and rice (green). (B-D) GO terms in onion (brown) and rice (green) associated with; (B) cellular component, (C) molecular function, and (D) biological process.

doi:10.1371/journal.pone.0166568.g002

Transcription factors in bulb onion

Transcription factors play an important role in plant development and stress responses [35]. A wide range of TFs have been identified and characterized in different plant species [36]. We identified 1837 bulb onion transcripts encoding orthologues of rice transcription factors and grouped into 55 families (Fig 4). The most highly represented transcription factor families were bHLH (162 transcripts), NAC (147 transcripts), EFR (132 transcripts), MYB (121 transcripts), WRKY (109 transcripts) and C2H2 (105 transcripts) (Fig 4). These transcription

Identification and expression analysis of male fertility genes

Male reproductive development and fertility are important agronomical traits in crop plants. The identification of genes involved in these processes allows better understanding of the molecular mechanism of male fertility [13] and will assist breeders to develop male sterile lines to utilize in heterosis breeding [13, 37]. Using our transcriptome data which was derived from developing flower buds and flowers (and other tissues) of normal male fertile plants, we found potential orthologues of a range of rice genes involved in male fertility and flower development (Table 3). These genes are also present in other plant species and have conserved functions indicating common mechanisms of flower development [38–39].

We identified a number of flower development genes, including the MADS box genes *PIS-TILLATA*, *AGAMOUS*, *SEPALLATA3*, *APETALLA3* and *AGAMOUS LIKE6*. These MADS box genes determined flower organ identity, and mutations in some of these genes can result in male sterility [66–69]. The expression pattern of floral meristem genes varies in different developmental stages across a wide range of plants [31, 70–71]. We found that *AGAMOUS*, *AGL6*, *AP3* and *SEPALLATA3* were expressed in bulb onion flowers (from unopened flowers to older fully open flowers) but not in immature flower heads (Fig 5). *PIS-TILLATA* had a similar expression pattern to *AP3* and the floral meristem identity genes, but was also detected at relatively high levels in bulb onion leaves (Fig 5). *PIS-TILLATA* has been found to be also expressed in the leaves and roots in different plants but their function in vegetative organs is still unknown [72–75]. The flower specific MADS genes we have identified in bulb onion could be mutated to generate male specific mutants. For example, the rice *AGAMOUS* (also known as *OsMADS3*) mutant plants show severe defects in stamen identity and lodicule number which leads to male sterility [76]. A naturally occurring mutation induced by retrotransposon insertion in *OsMADS3* has recently been identified, which causes recessive male-sterility but retains good agronomical performance, so it could be used as an elite line for recurrent selection [68].

The timely degradation of tapetal cells is a prerequisite for the development of viable pollen grains. *PERSISTANT TAPETAL CELL1* (*PTC1*) is a rice orthologue of Arabidopsis *MALE STERILITY1* (*MS1*) gene encoding a Plant Homeodomain (PHD) protein that regulates programmed tapetal development and pollen formation [77–79]. *MALE MEIOCYTE DEATH1* (*MMD1*) is another PHD protein involved in the regulation of gene expression during meiosis mutations [51]. Mutations in these genes results in complete male sterility in *Arabidopsis*, rice and barley [51, 77–80]. In the bulb onion transcriptome dataset we found two contigs having characteristic PHD domains encoding *PERSISTANT TAPETAL CELL1* (*PTC1*) and *MALE MEIOCYTE DEATH1* (*MMD1*). Other transcripts encoding proteins that are required for male fertility are listed in Table 3 (the sequences of onion contigs corresponding to these genes is given in S3 Table). As the molecular mechanism controlling floral development is largely conserved across plant species [38–39], some of the candidate male fertility genes we have identified in bulb onion could provide excellent targets for engineering new male sterile lines.

A new method of generating hybrid seed has been developed in maize [81]. This involves identifying or generating a male sterile mutant and then adding three transgenes to firstly complement the mutant to recover male fertile plants; second to prevent pollen formation so that the restoration of male fertility can only be maternally inherited, and thirdly to provide a easily detected fluorescent reporter protein ensuring any contaminating seed containing the transgenes is easily detected [81]. This system provides a simple way of generating male sterile female plants for hybrid seed production, however it is necessary to first generate a male sterile mutant. This has now been done using the CRISPR/Cas9 genome editing technique [82–83].

Table 3. Bulb onion orthologues of rice genes involved in male fertility and floral development identified using BLAST searches.

Gene Name	Bulb onion contigs ^a	Rice	% a.a Identity	Mutant	References
<i>CARBON SATRVED ANTHEAS (CSA)</i>	c244438_g1_i1	LOC_Os01g16810	51.70	Reduced levels of carbohydrates in anthers and it causes male sterility.	[40]
<i>PISTILLATA (OsMADS4)</i>	c95167_g1_i1	LOC_Os5g34940	55.50	Abnormal lodicule development	[41]
<i>AGAMOUS (OsMADS3)</i>	c34046_g1_i1	LOC_Os01g10504	60.40	Defective anther wall, aborted microspores	[42]
<i>SEPALLATA3 (OsMADS7)</i>	c216523_g1_i1	LOC_Os08g41950	53.50	homeotic changes of lodicules, stamens and carpels into palea/lemma-like organs, and a loss of floral determinacy.	[43]
<i>AGAMOUS LIKE 6 (OsMADS6)</i>	c129731_g1_i1	LOC_Os02g45770	70.90	Abnormal palea and lodicule development	[44]
<i>TGA9</i>	c151564_g1_i1	LOC_Os11g05480	66.40	Abnormal anther	[45]
<i>JAGGED (STAMENLESS)</i>	c30443_g1_i1	LOC_Os01g03840	48.40	Abnormal lemma and palea, Defective stamen	[46]
<i>MPK3</i>	c121397_g1_i1	LOC_Os03g17700	82.90	Defective anther	[47]
<i>MPK6</i>	c123481_g1_i1	LOC_Os06g06090	86.70	Defective anther	[47]
<i>SWI1 (OsAM1)</i>	c201101_g1_i1	LOC_Os03g44760	50.00	Arrested Meiosis at earlyt prophase 1	[48]
<i>TPD1 (TDL1A)</i>	c21970_g1_i1	LOC_Os12g28750	47.20	Multiple megaspore mother cells	[49]
<i>EMS1 (MSP1)</i>	c101027_g1_i1	LOC_Os01g68870	69.70	Excessive number of male and female sporocytes, lack of the tapetum	[50]
<i>MMD1</i>	c3619_g1_i1	LOC_Os03g50780	42.10	Male sterile due to a defect in meiosis II	[51]
<i>WAX-DEFICIENT ANTHEA 1</i>	c115430_g1_i1	LOC_Os10g33250	64.70	Defect in pollen exine formation	[52]
<i>REDUCED ADH ACTIVITY</i>	c78051_g1_i1	LOC_Os09g35000	50.70	-	-
<i>PERSISTENT TAPETAL CELL1</i>	c179910_g1_i1	LOC_Os09g27620	45.60	Mutants fail to make functional pollen; pollen degenerates after microspore release and the tapetum also appears abnormally vacuolated.	[53]
<i>AP3 SUPERWOMAN1 (OsMADS16)</i>	c229753_g1_i1	LOC_Os06g49840	67.80	transformation of the lodicule and the stamen into mrp- and carpel-like organs, respectively	[54]
<i>ENT-KAURENE OXIDASE</i>	c117603_g1_i2	LOC_Os06g02019	57.80	Defects in exine formation	[55]
<i>RAS-RELATED NUCLEAR PROTEIN GTPASE</i>	c112324_g1_i1	LOC_Os05g49890	92.30	-	-
<i>IMPORTIN BETA1</i>	c223739_g1_i1	LOC_Os05g28510	69.60	-	-
<i>ATSIZ1/SIZ1</i>	c119191_g13_i1	LOC_Os05g03430	75	-	-
<i>SOLO DANCERS</i>	c104859_g1_i1	LOC_Os03g12414	25.20	Defects in homolog interaction, bivalanet and meiotic crossover formation	[56]
<i>SUCROSE TRANSPORTER</i>	c87072_g1_i1	LOC_Os03g07480	66.80	-	-
<i>CYTOCHROME P-450B</i>	c74282_g1_i1	LOC_Os03g07250	68.10	Aborted pollen grains	[57]
<i>MYOSIN XI B</i>	c106734_g1_i1	LOC_Os02g57190	68.40	male sterility under short-day-length (SD) and fertility under long-day-length (LD) conditions	[58]
<i>APOPTOSIS INHIBITOR 5</i>	c81284_g1_i1	LOC_Os02g20930	66.80	Inhibition of tapetal PCD and aborted pollen	[59]
<i>ENT-KAURENE SYNTHASE</i>	c83938_g1_i2	LOC_Os02g17780	58.60	-	-
<i>SHOOTLESS2</i>	c78229_g1_i2	LOC_Os01g34350	59.30	Abnormal stamen development	[60]
<i>DMC1</i>	c43784_g1_i1	LOC_Os11g04954	87.10	Defects synapsis and crossing over	[61]
<i>ZEP1</i>	c116971_g1_i3	LOC_Os04g37960	44.60	Defects in synaptoneal complex assembly	[62]
<i>POLLEN SEMI-STERILITY1</i>	c115973_g1_i1	LOC_Os08g02380	40	Reduced pollen viability and anther dehiscence	[63]
<i>REC8</i>	c107330_g1_i3	LOC_Os01g67250	26.80	Defects in homologous chromosome pairing and telomere formation	[64]
<i>LIS1</i>	c65717_g1_i1	LOC_Os08g06480	81.20	Defects in male gametophyte and male sterility	[65]

^a The sequence of these contigs is provided in S3 Table.

doi:10.1371/journal.pone.0166568.t003

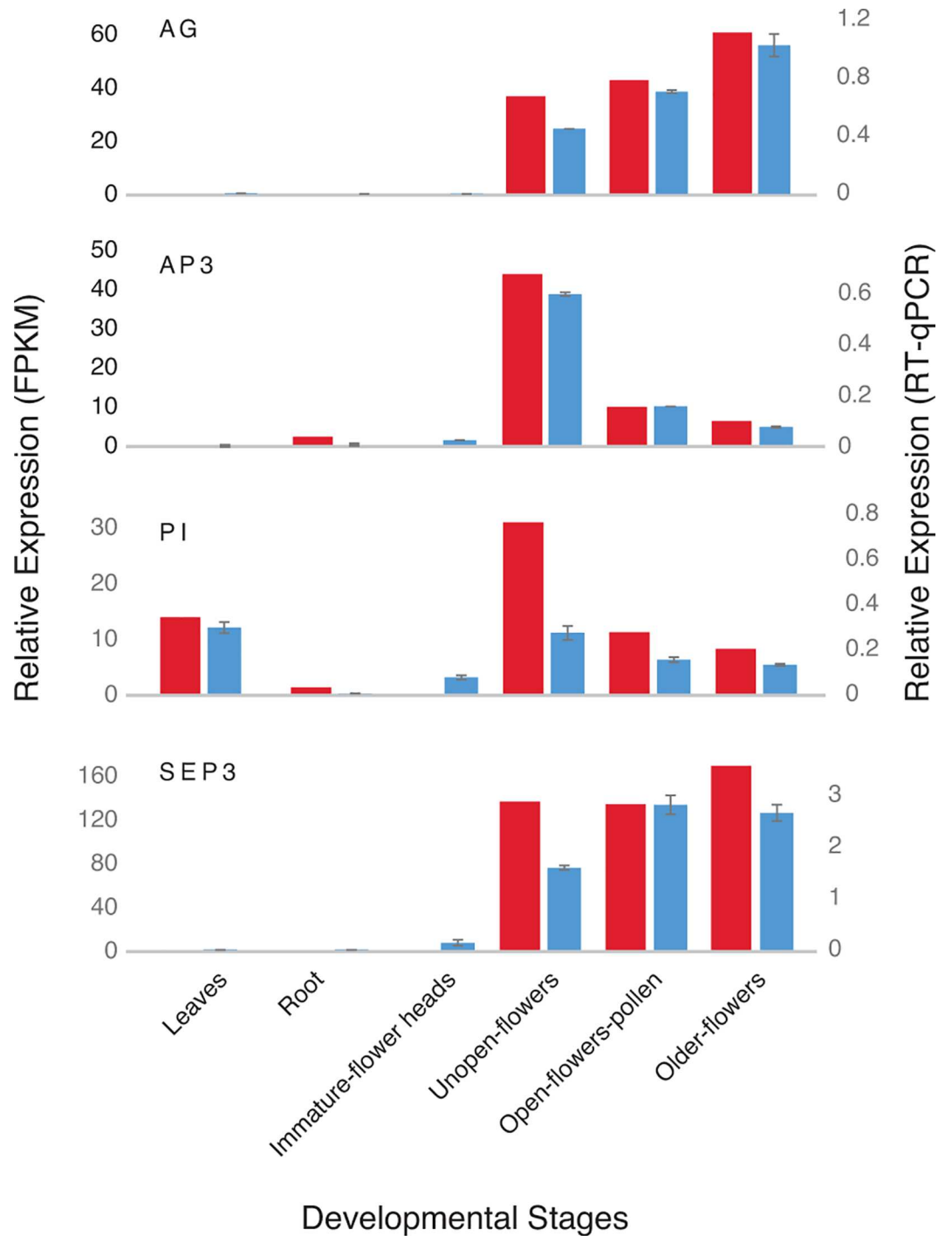


Fig 5. Relative expression of the floral meristem identity genes *AGAMOUS* (*AG*), *APETALLA3* (*AP3*), *PISTILLATA* (*P1*) and *SEPALLATA3* (*SEP3*) across different onion organs. Expression determined from the RNAseq data is shown in red and RT-PCR data is shown in blue, with data represented by an average \pm S.E. of three samples, with transcripts normalized to actin and β -tubulin.

doi:10.1371/journal.pone.0166568.g005

Conclusions

High heterozygosity and inbreeding depression hampers onion improvement and genetic programs but can be counteracted by using double haploid lines in genetic and molecular biology research projects. In this context we developed a transcriptome dataset using double haploid “CUDH2107” as reference line to provide more genomic resources for the *Allium* research community. The development of a transcriptome assembly from different development stages of bulb onion is a valuable genomic resource for better understanding the genetic and molecular basis of various traits. In the present investigation, a transcriptome dataset has been generated from different vegetative and reproductive organs. This dataset was explored to identify genes involved in male fertility and examine their expression in different organs. The next step would be to functionally characterize these genes to identify those that could be mutated to develop male sterile lines for hybrid production. A variety of approaches have been used for the production of a transgenic male sterility–fertility restoration system [37, 81–86]. Targeted mutagenesis has been utilized in maize to induce mutations in male fertility genes [81–83]. The use of genome editing techniques, such as CRISPR/cas9, provides a new way to induce specific mutations in genes regulating anther and pollen development. The ability to engineer sterility in bulb onion would remove the limitation of using a single source of male sterility (CMS-S), and could broaden the genetic base of F₁ hybrids.

Supporting Information

S1 Table. Description of samples used and their NCBI SRA links.
(DOCX)

S2 Table. List of primer Sequences used in present investigation.
(DOCX)

S3 Table. Sequences in FASTA of bulb onion orthologues of rice genes involved in male fertility and floral development identified using BLAST searches.
(XLSX)

Acknowledgments

Jiffinvir S. Khosa's PhD research is supported by an Indian Council of Agricultural Research (ICAR) International fellowship, an Otago University doctoral scholarship, and a New Zealand Onion Industry Postgraduate Scholarship. We thank Bronwyn Carlisle (University of Otago) for assisting with figure preparations.

Author Contributions

Conceptualization: JSK JM RCM.

Data curation: JSK JM.

Funding acquisition: JSK JM RCM.

Investigation: JSK RL MPJ JM.

Project administration: JM RCM.

Software: JSK SB JM.

Supervision: JL JM RCM.

Visualization: JSK SB JM RCM.

Writing – original draft: JSK JM RCM.

Writing – review & editing: JSK JL JM RCM.

References

1. McCallum J. Onion. In: Kole C, editor. *Genome Mapping and Molecular Breeding in Plants Volume 5*. Springer, Netherlands; 2007. pp. 331–342.
2. Khosa JS, McCallum J, Dhatt AS, Macknight RC. Enhancing onion breeding using molecular tools. *Plant Breed* 2016; 135: 9–20. doi: [10.1111/pbr.12330](https://doi.org/10.1111/pbr.12330)
3. Martin LB, Fei Z, Giovannoni JJ, Rose JK. Catalyzing plant science research with RNA-seq. *Front Plant Sci*. 2013; 4:66. doi: [10.3389/fpls.2013.00066](https://doi.org/10.3389/fpls.2013.00066) PMID: [23554602](https://pubmed.ncbi.nlm.nih.gov/23554602/)
4. Varshney RK, Terauchi R, McCouch SR. Harvesting the Promising Fruits of Genomics: Applying Genome Sequencing Technologies to Crop Breeding. *PLoS Biol*. 2014; 12(6):e1001883. doi: [10.1371/journal.pbio.1001883](https://doi.org/10.1371/journal.pbio.1001883) PMID: [24914810](https://pubmed.ncbi.nlm.nih.gov/24914810/)
5. Farrell JD, Byrne S, Paina C, Asp T. De Novo Assembly of the Perennial Ryegrass Transcriptome Using an RNA-Seq Strategy. *PLoS One*. 2014; 9(8):e103567. doi: [10.1371/journal.pone.0103567](https://doi.org/10.1371/journal.pone.0103567) PMID: [25126744](https://pubmed.ncbi.nlm.nih.gov/25126744/)
6. Zhang H, Tan E, Suzuki Y, Hirose Y, Kinoshita S, Okano H et al. et al. Dramatic improvement in genome assembly achieved using doubled-haploid genomes. *Sci Rep*. 2014; 4:6780. doi: [10.1038/srep06780](https://doi.org/10.1038/srep06780) PMID: [25345569](https://pubmed.ncbi.nlm.nih.gov/25345569/)
7. Alan AR, Brants A, Cobb E, Goldschmied P, Mutschler MA, Earle ED. Fecund gynogenic lines from onion (*Allium cepa* L.) breeding materials. *Plant Sci*. 2004; 167:1055–1066. doi: [10.1016/j.plantsci.2004.06.007](https://doi.org/10.1016/j.plantsci.2004.06.007)
8. Baldwin S, Revanna R, Thomson S, Pither-Joyce M, Wright K, Crowhurst R, et al. A Toolkit for bulk PCR-based marker design from next-generation sequence data: application for development of a framework linkage map in bulb onion (*Allium cepa* L.). *BMC Genomics*. 2012; 13:637. doi: [10.1186/1471-2164-13-637](https://doi.org/10.1186/1471-2164-13-637) PMID: [23157543](https://pubmed.ncbi.nlm.nih.gov/23157543/)
9. Lee R, Baldwin S, Kenel F, McCallum J, Macknight R. *FLOWERING LOCUS T* genes control onion bulb formation and flowering. *Nat Commun*. 2013; 4:2884. doi: [10.1038/ncomms3884](https://doi.org/10.1038/ncomms3884) PMID: [24300952](https://pubmed.ncbi.nlm.nih.gov/24300952/)
10. Duangjit J, Bohanec B, Chan AP, Town CD, Havey MJ. Transcriptome sequencing to produce SNP-based genetic maps of onion. *Theor Appl Genet*. 2013; 126(8):2093–2101. doi: [10.1007/s00122-013-2121-x](https://doi.org/10.1007/s00122-013-2121-x) PMID: [23689743](https://pubmed.ncbi.nlm.nih.gov/23689743/)
11. Hyde P, Earle E, Mutschler M. Doubled haploid onion (*Allium cepa* L.) lines and their impact on hybrid performance. *Hort Sci* 2012; 47: 1690–1695.
12. McCallum J, Baldwin S, Thomson S, Pither-Joyce M, Kenel F, Lee R, et al. Molecular genetics analysis of onion (*Allium cepa* L.) adaptive physiology of bulb. *Acta Hort*. 2014; 1110: 71–76.
13. Chen L, Liu YG. Male Sterility and Fertility Restoration in Crops. *Annu Rev Plant Biol*. 2014; 65:579–606. doi: [10.1146/annurev-arplant-050213-040119](https://doi.org/10.1146/annurev-arplant-050213-040119) PMID: [24313845](https://pubmed.ncbi.nlm.nih.gov/24313845/)
14. Yoshida H, Nagato Y. Flower development in rice. *J Exp Bot*. 2011; 62(14):4719–4730. doi: [10.1093/jxb/err272](https://doi.org/10.1093/jxb/err272) PMID: [21914655](https://pubmed.ncbi.nlm.nih.gov/21914655/)
15. Guo JX, Liu YG. Molecular control of male reproductive development and pollen fertility in rice. *J Integr Plant Biol*. 2012; 54(12):967–978. doi: [10.1111/j.1744-7909.2012.01172.x](https://doi.org/10.1111/j.1744-7909.2012.01172.x) PMID: [23025662](https://pubmed.ncbi.nlm.nih.gov/23025662/)
16. Kim B, Kim K, Yang TJ, Kim S. Completion of the mitochondrial genome sequence of onion (*Allium cepa* L.) containing the CMS-S male-sterile cytoplasm and identification of an independent event of the ccmF N gene split. *Current Genetics* doi: [10.1007/s00294-016-0595-1](https://doi.org/10.1007/s00294-016-0595-1) PMID: [27016941](https://pubmed.ncbi.nlm.nih.gov/27016941/)
17. Kim S, Kim CW, Park M, Choi D. Identification of candidate genes associated with fertility restoration of cytoplasmic male-sterility in onion (*Allium cepa* L.) using a combination of bulked segregant analysis and RNA-seq. *Theor Appl Genet*. 2015; 128(11):2289–2299. doi: [10.1007/s00122-015-2584-z](https://doi.org/10.1007/s00122-015-2584-z) PMID: [26215184](https://pubmed.ncbi.nlm.nih.gov/26215184/)
18. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol*. 2011; 29(7):644–652. doi: [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883) PMID: [21572440](https://pubmed.ncbi.nlm.nih.gov/21572440/)
19. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. 2013; 8(8):1494–1512. doi: [10.1038/nprot.2013.084](https://doi.org/10.1038/nprot.2013.084) PMID: [23845962](https://pubmed.ncbi.nlm.nih.gov/23845962/)

20. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV and Zdobnoy EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015; 31(19):3210–3212. doi: [10.1093/bioinformatics/btv351](https://doi.org/10.1093/bioinformatics/btv351) PMID: [26059717](https://pubmed.ncbi.nlm.nih.gov/26059717/)
21. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990; 215(3):403–410. doi: [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
22. Sun X, Zhou S, Meng F, Liu S. *De novo* assembly and characterization of the garlic (*Allium sativum*) bud transcriptome by Illumina sequencing. *Plant Cell Rep*. 2012; 31(10):1823–1828. doi: [10.1007/s00299-012-1295-z](https://doi.org/10.1007/s00299-012-1295-z) PMID: [22684307](https://pubmed.ncbi.nlm.nih.gov/22684307/)
23. Kim S, Kim MS, Kim YM, Yeom SI, Cheong K, Kim KT, et al. Integrative structural annotation of *de novo* RNA-Seq provides an accurate reference gene set of the enormous genome of the onion (*Allium cepa* L.). *DNA Res*. 2015; 22(1):19–27. doi: [10.1093/dnares/dsu035](https://doi.org/10.1093/dnares/dsu035) PMID: [25362073](https://pubmed.ncbi.nlm.nih.gov/25362073/)
24. Kamenetsky R, Faigenboim A, Mayer ES, Michael TB, Gershberg C, Kimhi S, et al. Integrated transcriptome catalogue and organ-specific profiling of gene expression in fertile garlic (*Allium sativum* L.). *BMC Genomics*. 2015; 16:12. doi: [10.1186/s12864-015-1212-2](https://doi.org/10.1186/s12864-015-1212-2) PMID: [25609311](https://pubmed.ncbi.nlm.nih.gov/25609311/)
25. Tsukazaki H, Yaguchi S, Sato S, Hirakawa H, Katayose Y, Kanamori H, et al. Development of transcriptome shotgun assembly-derived markers in bunching onion (*Allium fistulosum*). *Molecular Breeding*. 2015; 35:55. doi: [10.1007/s11032-015-0265-x](https://doi.org/10.1007/s11032-015-0265-x)
26. Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, et al. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice (N Y)*. 2013; 6(1):4. doi: [10.1186/1939-8433-6-4](https://doi.org/10.1186/1939-8433-6-4) PMID: [24280374](https://pubmed.ncbi.nlm.nih.gov/24280374/)
27. Min X.J, Butler G, Storms R, Tsang A. OrfPredictor: predicting protein-coding regions in EST-derived sequences. *Nucleic Acids Res*. 2005; 33(Web Server issue):W677–680. doi: [10.1093/nar/gki394](https://doi.org/10.1093/nar/gki394) PMID: [15980561](https://pubmed.ncbi.nlm.nih.gov/15980561/)
28. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without reference genome. *BMC Bioinformatics*. 2011; 12:323. doi: [10.1186/1471-2105-12-323](https://doi.org/10.1186/1471-2105-12-323) PMID: [21816040](https://pubmed.ncbi.nlm.nih.gov/21816040/)
29. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of differential gene expression data. *Bioinformatics*. 2010; 26(1):139–140. doi: [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616) PMID: [19910308](https://pubmed.ncbi.nlm.nih.gov/19910308/)
30. Fritsch RM, Friesen N. Evolution, Domestication and Taxonomy. In: Rabinowitch HD, Currah L, editors. *Allium Crop Science: Recent Advances*, CAB International, Wallingford, UK; 2002. pp. 5–30.
31. Janssen T, Bremer K. The age of major monocot groups inferred from 800+ rbcL sequences. *Botanical Journal of the Linnean Society*. 2004; 146: 385–398. doi: [10.1111/j.1095-8339.2004.00345.x](https://doi.org/10.1111/j.1095-8339.2004.00345.x)
32. Šmardaa P, Bureša P, Horováa L, Leitchb IJ, Mucinac L, Pacinie E, et al. Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proc Natl Acad Sci U S A*. 2014; 111(39): E4096–102. doi: [10.1073/pnas.1321152111](https://doi.org/10.1073/pnas.1321152111) PMID: [25225383](https://pubmed.ncbi.nlm.nih.gov/25225383/)
33. Kuhl J, Cheung F, Yuan Q, Martin W, Zewdie Y, McCallum J, et al. A unique set of 11,008 expressed sequence tags (EST) reveals expressed sequence and genomic differences between monocot order asparagales and poales. *Plant Cell*. 2004; 16(1):114–25. doi: [10.1105/tpc.017202](https://doi.org/10.1105/tpc.017202) PMID: [14671025](https://pubmed.ncbi.nlm.nih.gov/14671025/)
34. Glémin S, Clément Y, David J, Ressayre A. GC content evolution in coding regions of angiosperm genomes: a unifying hypothesis. *Trends Genet*. 2014; 30(7):263–270. doi: [10.1016/j.tig.2014.05.002](https://doi.org/10.1016/j.tig.2014.05.002) PMID: [24916172](https://pubmed.ncbi.nlm.nih.gov/24916172/)
35. Jin J, He K, Tang X, Li Z, Lv L, Zhao Y, et al. An Arabidopsis Transcriptional Regulatory Map Reveals Distinct Functional and Evolutionary Features of Novel Transcription Factors. *Mol Biol Evol*. 2015; 32(7):1767–1773. doi: [10.1093/molbev/msv058](https://doi.org/10.1093/molbev/msv058) PMID: [25750178](https://pubmed.ncbi.nlm.nih.gov/25750178/)
36. Hong JC. General Aspects of Plant Transcription Factor Families. In Gonzalez DH, editor. *Plant Transcription Factors: Evolutionary, Structural and Functional Aspects*. (Ed). Academic Press; 2016. pp. 35–56.
37. Wang K, Peng X, Ji Y, Yang P, Zhu Y, Li S. Gene, protein, and network of male sterility in rice. *Front Plant Sci*. 2013; 4:92. doi: [10.3389/fpls.2013.00092](https://doi.org/10.3389/fpls.2013.00092) PMID: [23596452](https://pubmed.ncbi.nlm.nih.gov/23596452/)
38. Wellmer F, Bowman JL, Davies B, Ferrándiz C, Fletcher JC, Franks RG, et al. Flower development: open questions and future directions. *Methods Mol Biol*. 2014; 1110:103–124. doi: [10.1007/978-1-4614-9408-9_5](https://doi.org/10.1007/978-1-4614-9408-9_5) PMID: [24395254](https://pubmed.ncbi.nlm.nih.gov/24395254/)
39. Fernández Gómez J, Talle B, Wilson ZA. Anther and pollen development: A conserved developmental pathway. *J Integr Plant Biol*. 2015; 57(11):876–891. doi: [10.1111/jipb.12425](https://doi.org/10.1111/jipb.12425) PMID: [26310290](https://pubmed.ncbi.nlm.nih.gov/26310290/)
40. Zhang H, Liang W, Yang X, Luo X, Jiang N, Ma H, Zhang D. Carbon Starved Anther Encodes a MYB Domain Protein That Regulates Sugar Partitioning Required for Rice Pollen Development. *The Plant Cell* 2010; 22(3): 672–689. doi: [10.1105/tpc.109.073668](https://doi.org/10.1105/tpc.109.073668) PMID: [20305120](https://pubmed.ncbi.nlm.nih.gov/20305120/)

41. Yao SG, Ohmori S, Kimizu M, Yoshida H. Unequal Genetic Redundancy of Rice PISTILLATA Orthologs, OsMADS2 and OsMADS4, in Lodicule and Stamen Development. *Plant Cell Physiol.* 2008; 49(5):853–857. doi: [10.1093/pcp/pcn050](https://doi.org/10.1093/pcp/pcn050) PMID: [18378529](https://pubmed.ncbi.nlm.nih.gov/18378529/)
42. Yamaguchi T, Lee DY, Miyao A, Hirochika H, An G, Hirano HY. Functional diversification of the two C-class MADS box genes OSMADS3 and OSMADS58 in *Oryza sativa*. *Plant Cell.* 2006; 18(1):15–28 doi: [10.1105/tpc.105.037200](https://doi.org/10.1105/tpc.105.037200) PMID: [16326928](https://pubmed.ncbi.nlm.nih.gov/16326928/)
43. Cui R, Han J, Zhao S, Su K, Wu F, Du X, et al. Functional conservation and diversification of class E floral homeotic genes in rice (*Oryza sativa*). *Plant J.* 2010; 61(5):767–781. doi: [10.1111/j.1365-313X.2009.04101.x](https://doi.org/10.1111/j.1365-313X.2009.04101.x) PMID: [20003164](https://pubmed.ncbi.nlm.nih.gov/20003164/)
44. Ohmori S, Kimizu M, Sugita M, Miyao A, Hirochika H, Uchida E et al. MOSAIC FLORAL ORGANS1, an AGL6-Like MADS Box Gene, Regulates Floral Organ Identity and Meristem Fate in Rice. *Plant Cell.* 2009; 21(10):3008–3025. doi: [10.1105/tpc.109.068742](https://doi.org/10.1105/tpc.109.068742) PMID: [19820190](https://pubmed.ncbi.nlm.nih.gov/19820190/)
45. Murmu J, Bush MJ, DeLong C, Li S, Xu M, Khan M, et al. Arabidopsis basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. *Plant Physiol.* 2010; 154(3):1492–1504. doi: [10.1104/pp.110.159111](https://doi.org/10.1104/pp.110.159111) PMID: [20805327](https://pubmed.ncbi.nlm.nih.gov/20805327/)
46. Xiao H, Tang J, Li Y, Wang W, Li X, Jin L, et al. STAMENLESS 1, encoding a single C2H2 zinc finger protein, regulates floral organ identity in rice. *Plant J.* 2009; 59(5):789–801. doi: [10.1111/j.1365-313X.2009.03913.x](https://doi.org/10.1111/j.1365-313X.2009.03913.x) PMID: [19453444](https://pubmed.ncbi.nlm.nih.gov/19453444/)
47. Hord CL, Sun Y-J, Pillitterie L, Toriie KU, Wang H, Zhang S et al. Regulation of Arabidopsis early anther development by the mitogen-activated protein kinases, MPK3 and MPK6, and the ERECTA and related receptor-like kinases. *Mol Plant.* 2008; 1(4):645–658. doi: [10.1093/mp/ssn029](https://doi.org/10.1093/mp/ssn029) PMID: [19825569](https://pubmed.ncbi.nlm.nih.gov/19825569/)
48. Che L, Tang D, Wang K, Wang M, Zhu K, Yu H, et al. OsAM1 is required for leptotene-zygotene transition in rice. *Cell Res.* 2011; 21(4):654–665. doi: [10.1038/cr.2011.7](https://doi.org/10.1038/cr.2011.7) PMID: [21221128](https://pubmed.ncbi.nlm.nih.gov/21221128/)
49. Zhao X1, de Palma J, Oane R, Gamuyao R, Luo M, Chaudhury A, et al. OsTDL1A binds to the LRR domain of rice receptor kinase MSP1, and is required to limit sporocyte numbers. *Plant J.* 2008; 54(3):375–387. doi: [10.1111/j.1365-313X.2008.03426.x](https://doi.org/10.1111/j.1365-313X.2008.03426.x) PMID: [18248596](https://pubmed.ncbi.nlm.nih.gov/18248596/)
50. Nonomura K, Miyoshi K, Eiguchi M, Suzuki T, Miyao A, Hirochika H et al. The MSP1 gene is necessary to restrict the number of cells entering into male and female sporogenesis and to initiate anther wall formation in rice. *Plant Cell.* 2003; 15(8):1728–1739. doi: [10.1105/tpc.012401](https://doi.org/10.1105/tpc.012401) PMID: [12897248](https://pubmed.ncbi.nlm.nih.gov/12897248/)
51. Yang X, Makaroff CA and Ma H. The Arabidopsis MALE MEIOCYTE DEATH1 gene encodes a PHD-finger protein that is required for male meiosis. *Plant Cell.* 2003; 15(6):1281–95. doi: [10.1105/tpc.010447](https://doi.org/10.1105/tpc.010447) PMID: [12782723](https://pubmed.ncbi.nlm.nih.gov/12782723/)
52. Jung KH, Han MJ, Lee DY, Lee YS, Schreiber L, Franke R, et al. Wax-deficient anther1 Is Involved in Cuticle and Wax Production in Rice Anther Walls and Is Required for Pollen Development. *Plant Cell.* 2006; 18(11):3015–3032. doi: [10.1105/tpc.106.042044](https://doi.org/10.1105/tpc.106.042044) PMID: [17138699](https://pubmed.ncbi.nlm.nih.gov/17138699/)
53. Li H, Yuan Z, Vizcay-Barrena G, Yang C, Liang W, Zong J, Wilson ZA, Zhang D. PERSISTENT TAPETAL CELL1 encodes a PHD-finger protein that is required for tapetal cell death and pollen development in rice. *Plant Physiol.* 2011a; 156(2):615–630. doi: [10.1104/pp.111.175760](https://doi.org/10.1104/pp.111.175760) PMID: [21515697](https://pubmed.ncbi.nlm.nih.gov/21515697/)
54. Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H et al. SUPERWOMAN1 and DROOPING LEAF genes control floral organ identity in rice. *Development.* 2003; 130(4):705–718. doi.org/10.1105/tpc.018044. PMID: [12506001](https://pubmed.ncbi.nlm.nih.gov/12506001/)
55. Aya K, Hiwatashi Y, Kojima M, Sakakibara H, Ueguchi-Tanaka M, Hasebe M et al. The Gibberellin perception system evolved to regulate a pre-existing GAMYB-mediated system during land plant evolution. *Nat Commun.* 2011; 2:544. doi: [10.1038/ncomms1552](https://doi.org/10.1038/ncomms1552) PMID: [22109518](https://pubmed.ncbi.nlm.nih.gov/22109518/)
56. Chang L, Ma H, Xue HW. Functional conservation of the meiotic genes SDS and RCK in male meiosis in the monocot rice. *Cell Res.* 2009; 19(6):768–782. doi: [10.1038/cr.2009.52](https://doi.org/10.1038/cr.2009.52) PMID: [19417775](https://pubmed.ncbi.nlm.nih.gov/19417775/)
57. Li H, Pinot F, Sauveplane V, Werck-Reichhart D, Diehl P, Schreiber L, et al. Cytochrome P450 Family Member CYP704B2 Catalyzes the ω -Hydroxylation of Fatty Acids and Is Required for Anther Cutin Biosynthesis and Pollen Exine Formation in Rice. *Plant Cell.* 2010; 22(1):173–190. doi: [10.1105/tpc.109.070326](https://doi.org/10.1105/tpc.109.070326) PMID: [20086189](https://pubmed.ncbi.nlm.nih.gov/20086189/)
58. Jiang SY, Cai M, Ramachandran S. *ORYZA SATIVA MYOSIN XI B* controls pollen development by photoperiod-sensitive protein localizations. *Dev Biol.* 2007; 304(2):579–592. doi: [10.1016/j.ydbio.2007.01.008](https://doi.org/10.1016/j.ydbio.2007.01.008) PMID: [17289016](https://pubmed.ncbi.nlm.nih.gov/17289016/)
59. Li X, Gao X, Wei Y, Deng L, Ouyang Y, Chen G, et al. Rice APOPTOSIS INHIBITOR5 coupled with two DEAD-box adenosine 5'-triphosphate-dependent RNA helicases regulates tapetum degeneration. *Plant Cell.* 2011b; 23(4):1416–1434. doi: [10.1105/tpc.110.082636](https://doi.org/10.1105/tpc.110.082636) PMID: [21467577](https://pubmed.ncbi.nlm.nih.gov/21467577/)

60. Toriba T, Suzaki T, Yamaguchi T, Ohmori Y, Tsukaya H, Hirano HY. Distinct regulation of adaxial-abaxial polarity in anther patterning in rice. *Plant Cell*. 2010; 22(5):1452–1462. doi: [10.1105/tpc.110.075291](https://doi.org/10.1105/tpc.110.075291) PMID: [20511295](https://pubmed.ncbi.nlm.nih.gov/20511295/)
61. Wang H, Hu Q, Tang D, Liu X, Du G, Shen Y, Li Y, Cheng Z. OsDMC1 Is Not Required for Homologous Pairing in Rice Meiosis. *Plant Physiol*. 2016; 171(1):230–241. doi: [10.1104/pp.16.00167](https://doi.org/10.1104/pp.16.00167) PMID: [26960731](https://pubmed.ncbi.nlm.nih.gov/26960731/)
62. Wang M, Kejian Wang K, Ding Tang D, Cunxu Wei C, Ming Li M, et al. The Central Element Protein ZEP1 of the Synaptonemal Complex Regulates the Number of Crossovers during Meiosis in Rice. *Plant Cell*. 2010; 22(2):417–430. doi: [10.1105/tpc.109.070789](https://doi.org/10.1105/tpc.109.070789) PMID: [20154151](https://pubmed.ncbi.nlm.nih.gov/20154151/)
63. Zhou S, Wang Y, Li W, Zhao Z, Ren Y, Wang Y, et al. Pollen Semi-Sterility1 Encodes a Kinesin-1–Like Protein Important for Male Meiosis, Anther Dehiscence, and Fertility in Rice. *Plant Cell*. 2011; 23(1):111–129. doi: [10.1105/tpc.109.073692](https://doi.org/10.1105/tpc.109.073692) PMID: [21282525](https://pubmed.ncbi.nlm.nih.gov/21282525/)
64. Shao T, Tang D, Wang K, Wang M, Che L, Qin B, et al. OsREC8 is essential for chromatid cohesion and metaphase I monopolar orientation in rice meiosis. *Plant Physiol*. 2011; 156(3):1386–1396. doi: [10.1104/pp.111.177428](https://doi.org/10.1104/pp.111.177428) PMID: [21606318](https://pubmed.ncbi.nlm.nih.gov/21606318/)
65. Gao X, Chen Z, Zhang J, Li X, Chen G, Li X, Wu C. OsLIS-L1 encoding a lissencephaly type-1-like protein with WD40 repeats is required for plant height and male gametophyte formation in rice. *Planta*. 2012; 235(4):713–727. doi: [10.1007/s00425-011-1532-7](https://doi.org/10.1007/s00425-011-1532-7) PMID: [22020753](https://pubmed.ncbi.nlm.nih.gov/22020753/)
66. Huang F, Xu G, Chi Y, Liu H, Xue Q, Zhao T, et al. A soybean MADS-box protein modulates floral organ numbers, petal identity and sterility. *BMC Plant Biol*. 2014; 14:89. doi: [10.1186/1471-2229-14-89](https://doi.org/10.1186/1471-2229-14-89) PMID: [24693922](https://pubmed.ncbi.nlm.nih.gov/24693922/)
67. Tsai WC, Pan ZJ, Hsiao YY, Chen LJ, Zhong-Jian Liu ZJ. Evolution and function of MADS-box genes involved in orchid floral development. *J Systematics Evolution*. 2014; 52(4):397–410. doi: [10.1111/jse.12010](https://doi.org/10.1111/jse.12010)
68. Zhang L, Mao D, Xing F, Bai X, Zhao H, Yao W, et al. Loss of function of OsMADS3 via the insertion of a novel retrotransposon leads to recessive male sterility in rice (*Oryza sativa*). *Plant Sci*. 2015; 238:188–197. doi: [10.1016/j.plantsci.2015.06.007](https://doi.org/10.1016/j.plantsci.2015.06.007) PMID: [26259187](https://pubmed.ncbi.nlm.nih.gov/26259187/)
69. Ai Y, Zhang Q, Wang W, Zhang C, Cao Z, Bao M, et al. Transcriptomic Analysis of Differentially Expressed Genes during Flower Organ Development in Genetic Male Sterile and Male Fertile *Tagetes erecta* by Digital Gene-Expression Profiling. *PLoS One*. 2016; 11(3):e0150892. doi: [10.1371/journal.pone.0150892](https://doi.org/10.1371/journal.pone.0150892) PMID: [26939127](https://pubmed.ncbi.nlm.nih.gov/26939127/)
70. Taoka K, Ohki I, Tsuji H, and Kojima C, Shimamoto K. Structure and function of florigen and the receptor complex. *Trends Plant Sci*. 2013; 18(5):287–294. doi: [10.1016/j.tplants.2013.02.002](https://doi.org/10.1016/j.tplants.2013.02.002) PMID: [23477923](https://pubmed.ncbi.nlm.nih.gov/23477923/)
71. Stewart D, Graciet E, Wellmer F. Molecular and regulatory mechanisms controlling floral organ development. *FEBS J*. 2016; 283(10):1823–1830. doi: [10.1111/febs.13640](https://doi.org/10.1111/febs.13640) PMID: [26725470](https://pubmed.ncbi.nlm.nih.gov/26725470/)
72. Skipper M. Genes from the APETALA3 and PISTILLATA lineages are expressed in developing vascular bundles of the tuberous rhizome, flowering stem and flower Primordia of *Eranthis hyemalis*. *Ann Bot*. 2002; 89(1):83–88. doi: [10.1093/aob/mcf009](https://doi.org/10.1093/aob/mcf009) PMID: [12096822](https://pubmed.ncbi.nlm.nih.gov/12096822/)
73. Berbel A, Navarro C, Ferrándiz C, Cañas LA, Beltrán JP, Madueño F. Functional conservation of PISTILLATA activity in a pea homolog lacking the PI motif. *Plant Physiol*. 2005; 139(1):174–185. doi: [10.1104/pp.104.057687](https://doi.org/10.1104/pp.104.057687) PMID: [16113230](https://pubmed.ncbi.nlm.nih.gov/16113230/)
74. Poupin MJ, Federici F, Medina C, Matus JT, Timmermann T, Arce-Johnson P. Isolation of the three grape sub-lineages of B-class MADS box TM6, PISTILLATA and APETALA3 genes which are differentially expressed during flower and fruit development. *Gene*. 2007 Dec 1; 404(1–2):10–24 doi: [10.1016/j.gene.2007.08.005](https://doi.org/10.1016/j.gene.2007.08.005) PMID: [17920788](https://pubmed.ncbi.nlm.nih.gov/17920788/)
75. Lü S, Yinglun Fan Y, Like Liu L, Shujun Liu S, Wenhui Zhang W, Meng Z. Ectopic expression of TrPI, a *Taihania rupestris* (Rosaceae) PI ortholog, causes modifications of vegetative architecture in *Arabidopsis*. *J Plant Physiol*. 2010; 167(18):1613–1621. doi: [10.1016/j.jplph.2010.06.028](https://doi.org/10.1016/j.jplph.2010.06.028) PMID: [20828868](https://pubmed.ncbi.nlm.nih.gov/20828868/)
76. Yamaguchi T, Lee DY, Miyao A, Hirochika H, An G, Hirano HY. Functional diversification of the two C-class MADS box genes OSMADS3 and OSMADS58 in *Oryza sativa*. *Plant Cell*. 2006; 18(1):15–28 doi: [10.1105/tpc.105.037200](https://doi.org/10.1105/tpc.105.037200) PMID: [16326928](https://pubmed.ncbi.nlm.nih.gov/16326928/)
77. Ito T, Shinozaki K. The *MALE STERILITY1* gene of *Arabidopsis*, encoding a nuclear protein with a PHD-finger motif, is expressed in tapetal cells and is required for pollen maturation. *Plant Cell Physiol*. 2002; 43(11):1285–1292. doi: [10.1093/pcp/pcf154](https://doi.org/10.1093/pcp/pcf154) PMID: [12461128](https://pubmed.ncbi.nlm.nih.gov/12461128/)
78. Itoa T, Nagata N, Yoshiba Y, Ohme-Takagi M, Ma H, Shinozaki K. *Arabidopsis MALE STERILITY1* Encodes a PHD-Type Transcription Factor and Regulates Pollen and Tapetum Development. *Plant Cell*. 2007; 19(11):3549–3562. doi: [10.1105/tpc.107.054536](https://doi.org/10.1105/tpc.107.054536) PMID: [18032630](https://pubmed.ncbi.nlm.nih.gov/18032630/)

79. Li H, Yuan Z, Vizcay-Barrena G, Yang C, Liang W, Zong J, Wilson ZA, Zhang D. PERSISTENT TAPETAL CELL1 encodes a PHD-finger protein that is required for tapetal cell death and pollen development in rice. *Plant Physiol.* 2011a; 156(2):615–630. doi: [10.1104/pp.111.175760](https://doi.org/10.1104/pp.111.175760) PMID: [21515697](https://pubmed.ncbi.nlm.nih.gov/21515697/)
80. Fernández Gómez J, Wilson ZA. A barley PHD finger transcription factor that confers male sterility by affecting tapetal development. *Plant Biotechnol J.* 2014; 12(6):765–777. doi: [10.1111/pbi.12181](https://doi.org/10.1111/pbi.12181) PMID: [24684666](https://pubmed.ncbi.nlm.nih.gov/24684666/)
81. Wu Y, Fox TW, Trimnell MR, Wang L, Xu RJ, Cigan AM, Huffman GA, Garnaat CW, Hershey H, Albertsen MC. Development of a novel recessive genetic male sterility system for hybrid seed production in maize and other cross-pollinating crops. *Plant Biotechnol J.* 2016; 14(3):1046–54. doi: [10.1111/pbi.12477](https://doi.org/10.1111/pbi.12477) PMID: [26442654](https://pubmed.ncbi.nlm.nih.gov/26442654/)
82. Djukanovic V, Smith J, Lowe K, Yang M, Gao H, Jones S, et al. Male-sterile maize plants produced by targeted mutagenesis of the cytochrome P450-like gene (MS26) using a re-designed I-CreI homing endonuclease. *Plant J.* 2013; 76(5):888–899. doi: [10.1111/tpj.12335](https://doi.org/10.1111/tpj.12335) PMID: [24112765](https://pubmed.ncbi.nlm.nih.gov/24112765/)
83. Svitashv S, Young JK, Schwartz C, Gao H, Falco SC, Cigan AM. Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol.* 2015; 169(2):931–945. doi: [10.1104/pp.15.00793](https://doi.org/10.1104/pp.15.00793) PMID: [26269544](https://pubmed.ncbi.nlm.nih.gov/26269544/)
84. Singh SP, Singh SP, Pandey T, Singh RR, Sawant SV. A novel male sterility-fertility restoration system in plants for hybrid seed production. *Sci Rep.* 2015; 5:11274. doi: [10.1038/srep11274](https://doi.org/10.1038/srep11274) PMID: [26073981](https://pubmed.ncbi.nlm.nih.gov/26073981/)
85. Bohra A, Jha UC, Adhimoolam P, Bisht D, Singh NP. Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. *Plant Cell Rep.* 2016; 35(5):967–993. doi: [10.1007/s00299-016-1949-3](https://doi.org/10.1007/s00299-016-1949-3) PMID: [26905724](https://pubmed.ncbi.nlm.nih.gov/26905724/)
86. Li Q, Zhang D, Chen M, Liang W, Wei J, Qi Y, Yuan Z. (2016) Development of *japonica* photo-sensitive genic male sterile rice lines by editing *carbon starved anther* using CRISPR/Cas9. *J Genet Genomics.* 2016 43(6):415–9. doi: [10.1016/j.jgg.2016.04.011](https://doi.org/10.1016/j.jgg.2016.04.011). Epub 2016 May 4. PMID: [27317309](https://pubmed.ncbi.nlm.nih.gov/27317309/)