Co-expression network analysis of the genes associated with pistillody-stamen development in wheat

CURRENT STATUS: POSTED

Mingli Liao
China West Normal University

Zhenyong Chen
China West Normal University

Zaijun Yang
China West Normal University

Weiying Chen
China West Normal University

Shuhong Wei
China West Normal University

Zhengsong Peng
China West Normal University

DOI: 10.21203/rs.3.rs-16910/v1

SUBJECT AREAS
Epigenetics & Genomics

KEYWORDS
Co-expression network analysis, Pistillody-stamen development, Wheat, Transcription factor
Abstract
Background: Crop male sterility has great values in both theoretical research and breeding application. Wheat pistillody-stamen is an important male sterility phenomenon, and HTS-1 is an important pistillody-stamen material. However the molecular mechanism of HTS-1 stamens transformed into pistils or pistil-like structures remains a mystery. Weighted gene co-expression network analysis (WGCNA) are widely used to explore hub genes and gene interaction networks from high throughput data in various plants.

Results: In the present study, for exploring gene networks associated with wheat pistillody-stamen development, WGCNA was employed to analyze 11 RNA-sequencing (RNA-seq) data of wheat tissues, including stamens of CSTP, pistils and pistillody-stamen of HTS-1. 19 out of 25 merged modules were highly associated with specific wheat tissues, and the MEdarkseagreen1 module was highly related to wheat pistillody-stamen (correlation with weight \( r = 0.7 \), correlation p-value \( p = 0.02 \)). Then 180 genes about wheat flower development were identified from the MEdarkseagreen1 module by GO term analysis. Among 180 genes, the hub gene number associated with anther, filament, style, and ovary development were 12, 3, 3, and 10, respectively. We compared the published pistillody related proteins with proteins of HTS-1 by BLAST. A total of 58 pistillody-stamen development associated proteins were validated by BLAST. MADS-box and YABBY transcription factor about pistillody-stamen development were also analyzed in wheat flower. There were 47 of MADS-box and 17 of YABBY transcription factors were identified. BLAST program was used to align the published pistillody associated MADS-box and YABBY transcription factors with transcription factors identified in wheat flower. Totally, 36 of 47 MADS-box and 14 of 17 YABBY transcription factors were considered to regulate the development of pistillody-stamen, which had never been reported yet.

Conclusion: These results have systematically identified the key candidate genes about the development of HTS-1 substructures flower. The tissue-specific correlation network analyses provide important insights into the molecular interactions underlying pistillody-stamen development.

Background
Common wheat (Triticum aestivum L.) is a major source of protein, vitamins and minerals for humans
According to the report of United States department of agriculture, the global output of wheat in 2017, 2018, and July 2019 were accounted for 761.88, 730.90, and 771.46 million tons, respectively [2]. Although the global output of wheat are steadily increased in recent years, some studies show that wheat yields are hard to meet demand of increasing human population [3]. Thus many researchers strive to improve grain productivity and qualities. The development of wheat flowers affect grain yields and qualities. There are various wheat flower mutant materials have been reported in the past decade. Mutant (cr)-N26 and (cr)-CSdt7BS are male sterility material, whose stamen of floret will transform into pistil or pistil-like structure (called pistillody) [4–6]. Both (cr)-N26 and (cr)-CSdt7BS mutants contain exogenous genes of goat weed, so the pistillody-stamen phenomena of (cr)-N26 and (cr)-CSdt7BS are caused by heterologous cytoplasmic male sterility [5]. In 2013, Peng et al. found a new pistillody mutant called HTS-1 from three pistil (TP) wheat [7, 8]. The genetic base of HTS-1 mutant was different from the previously reported mutant (cr)-N26 and (cr)-CSdt7BS [6, 8]. Since HTS-1 mutants do not contain exogenous cytoplasm, the pistillody-stamen phenomena of HTS-1 mutant is controlled by nuclear genes [8]. Previous reports show that the molecular mechanisms of wheat stamens transformed into pistils of floret are very complex. In 2002, Murai et al. found that reduced WAP3 expression could be associated with the induction of pistillody in (cr)-CSdt7BS [5]. The mitochondrial gene orf25 of Ae. crassa cytoplasm had been considered to be a candidate as the cytoplasmic factor related to pistillody in the alloplasmic wheat [9]. Zhu et al. reported a mitochondrial gene called orf260cra, which causing the stamen of (cr)-CSdt7BS transformed into pistillody [10]. Situ hybridization analysis showed that a calmodulin-binding protein gene called WCBP1 was highly expressed in the pistil-like stamens at early to late developmental stages. This suggested that WCBP1 gene played an important role in pistil-like stamen development [11]. The homologous transformation of stamens into pistils of HTS – 1 floret was determined by the interaction of two recessive karyogenes hts1 and hts2 [8]. In 2015, Yang et al. identified 206 differentially expressed genes (DEG) correlated to HTS-1 stamen and pistil development using RNA-seq analysis. 123 of 206 genes were down-regulated genes and the remaining of 83 genes were up-regulated genes [12]. In 2019, sun et al. reported a TaEPFL1 gene, which was highly expressed in pistillody-
stamens than in pistils and stamens. Heterologous expression of the TaEPFL1 gene in Arabidopsis caused shortened filaments and pedicels, which might induce stamens transformation into pistils or pistil-like structures in wheat [13]. Furthermore, some transcription factors of MADS-box and DROOPING LEAF gene of YABBY family have been considered to regulate in the development of wheat pistillody-stamen [6, 9, 14-18]. The expression pattern analysis indicate that TaDL1, TaDL2, and TaDL3 genes might cause the stamen of HTS-1 complete or partial transforming into pistils [16]. These researches let us to know the development of wheat pistillody stamen. Whereas the genes mechanisms of stamen transformation into pistil are still unclear.

With rapid improvement of high throughput sequencing and computer analysis technology, we can systematically study complicated biological problems using variety of high throughput sequencing data. Various network analysis methods have been constructed and widely used to reveal the interaction of molecular in high throughput sequencing data analysis, such as neural network analysis [19], protein-protein interaction network analysis [20] and WGCNA [21]. WGCNA is an important method for studying gene regulation network. According to gene expression levels in different tissues, genes with similar expression patterns can be clustered into the same module by WGCNA [21]. Therefore, WGCNA is widely used to explore hub genes, relationship between modules and traits, and gene interaction networks of flower development in various plants [22-24]. Particularly, Ramírez-González et al. using 209 RNA-seq samples from 22 tissue types analyzed gene expression pattern of wheat Azhurnaya cultivar, and established tissue- and stress-specific co-expression networks that revealed extensive coordination of homoeolog expression throughout wheat development [1]. These networks, including spikelet development networks, and detailed gene expression atlases were deposited in wheat eFP browser (http://bar.utoronto.ca) [1]. However networks in wheat eFP browser don't reveal and provide the information about mechanism of genes for pistillody-stamen development. Because Azhurnaya cultivar is not a pistillody-stamen wheat. Furthermore, systematic identifications of genes associated pistillody-stamen development by WGCNA have never been reported yet.

In the present study, to explore genes related to wheat pistillody-stamen development and their
interaction mechanisms, we employed WGCNA to analyze 11 RNA-sequencing (RNA-seq) data of wheat tissues, including stamens of CSTP, pistils and pistillody-stamen of HTS-1 [12]. We surveyed the relationship between modules and specific flower tissue from wheat RNA-seq data. Then we identified genes associated with pistillody-stamen development from the modules highly by KOBAS. In order to validate the results explored by WGCNA, expression pattern of 20 genes related to the development of pistillody-stamen in HTS-1 were performed by Quantitative real-time PCR (qRT-PCR). This work will provide further insights into the molecular mechanisms underlying wheat pistillody-stamen development.

Results

A total of 58,576 genes obtained by Salmon were used to construct WGCNA networks. There were 47,492 genes remained after filtering out the genes with low expression level (TPM < 1). Once the co-expression network was constructed, the next step was module identification. Modules in co-expression network were clusters of highly interconnected genes. Figure 1B showed the gene dendrogram clustering on Tom-based dissimilarity and modules obtained by the dynamic tree cut. There were 224 modules were clustered by the similarity of their eigengenes, and each module described by a branch of different color. If a module was highly correlated to other modules, they should be merged into one module. Finally, Fig. 1B depicted the major tree branches constitute 25 merged modules labeled by different colors (cutHeight = 0.25).

WGCNA enabled us to reveal the relationships between the modules and tissue types (Fig. 2). When the p-value was less than 0.05, it indicated that this module was significant correlation with the tissue type. Figure 2 depicted that 19 out of 25 modules were highly associated with specific wheat tissue types, excluding MEburlywood, MEdarksalmon, MEcoral2, MEcoral4, MEGrey, and MElavenderblush. Each tissues had at least associated with a specific modules, excluding pistil and late ovary. Figure 2 displayed that the MEdarkseagreen1 module was highly related to wheat pistillody-stamen (correlation with weight r = 0.7, correlation p-value p = 0.02). The MEmediumpurple3 (r = 0.67, p = 0.02) and MEturquoise (r = 0.99, p = 3e-09) modules were significantly correlated with wheat stamen. The genes of the MEcoral1 (r = 0.66, p = 0.03), MElightblue4 (r = 0.88, p = 4e-04), and
MElightskyblue2 (r = 0.84, p = 0.001) modules were massively expressed in booting spike. The MEchocolate4 module exhibited a significantly correlation with early ovary (r = 0.62, p = 0.04), which demonstrated that genes clustering in MEchocolate4 might be strongly affected with wheat early ovary development. The genes of MEsalmon1(r = 0.62, p = 0.04) and MEdarkgoldenrod1 module (r = 0.87, p = 5e-04) were highly associated with rachis, indicating that this group of genes might be regulating to wheat rachis development. The MEdeeppink2 (r = 0.97, p = 7e-07) and MEmediumpurple4 (r = 0.61, p = 0.05), and MEplum (r = 0.68, p = 0.02) modules were highly correlated with ear spike development.

To investigate genes related to HTS-1 flower development in MEdarkseagreen1 module, we performed a Gene Ontology (GO) analysis by Kobas 3.0 (Additional file 1: Table S1). In Additional file 1: Table S1, there were 691 GO term records related to wheat flower substructure development. Further analysis revealed that the 691 GO terms were identified from 180 genes (Additional file 1: Table S1). Table 1 depicted genes number corresponding to different GO terms of wheat flower development, which were accounting from Additional file 1: Table S1. Table 1 showed that there were 51 genes about anther development, which were assigned to 8 different GO terms. There were 5 genes related to wheat style development in GO term of GO:0048479. There were 3 wheat filament development associated genes in GO term of GO:0080086. A total of 40 genes related to ovary development were identified in GO term of GO:0035670, and there were 20 genes about ovule development in GO term of GO:0048481. Wheat stamen and pistil had 3 and 2 types of GO terms, respectively. Moreover, there were 423 genes involved in wheat flower development, which were belonging to 22 types of GO terms (Table 1).

| Tissue      | GO term       | Description                      | Number of Genes | Tissue      | GO term       | Description                      | Number of Genes |
|-------------|---------------|----------------------------------|-----------------|-------------|---------------|----------------------------------|-----------------|
| Anther      | GO:0009901    | anther dehiscence                | 3               | Flower      | GO:0048449    | floral organ development         | 12              |
| Anther      | GO:0048653    | anther development               | 27              | Flower      | GO:0048444    | floral organ morphogenesis       | 13              |
| Anther      | GO:0048654    | anther morphogenesis             | 4               | Flower      | GO:0048438    | floral whorl development         | 74              |
| Anther      | GO:0048657    | anther wall, tapetum, cell       | 4               | Flower      | GO:0009908    | flower development               | 166             |
| GO:0010234 | anther wall tapetum cell fate specification | Flower GO:0048574 | long-day photoperiodism, flowering | 5 |
|------------|--------------------------------------------|--------------------|----------------------------------|---|
| GO:0048658 | anther wall tapetum development | Flower GO:0010076 | maintenance of floral meristem identity | 1 |
| GO:0048656 | anther wall tapetum formation | Flower GO:0048497 | maintenance of floral organ identity | 9 |
| GO:0048655 | anther wall tapetum morphogenesis | Flower GO:0009910 | negative regulation of flower development | 5 |
| GO:0048479 | style development | Flower GO:0048579 | negative regulation of long-day photoperiodism, flowering | 4 |
| GO:0080086 | stamen filament development | Flower GO:0048577 | negative regulation of short-day photoperiodism, flowering | 3 |
| GO:0035670 | plant-type ovary development | Flower GO:0048573 | photoperiodism, flowering | 18 |
| GO:0048481 | plant ovule development | Flower GO:0009911 | positive regulation of flower development | 22 |
| GO:0048443 | stamen development | Flower GO:0009909 | regulation of flower development | 40 |
| GO:0048455 | stamen formation | Flower GO:0048586 | regulation of long-day photoperiodism, flowering | 4 |
| GO:0048448 | stamen morphogenesis | Flower GO:2000028 | regulation of photoperiodism, flowering | 8 |
| GO:0048440 | carpel development | Flower GO:0048587 | regulation of short-day photoperiodism, flowering | 5 |
| GO:0009875 | pollen-pistil interaction | Flower GO:0048575 | short-day photoperiodism, flowering | 6 |
| GO:0010582 | floral meristem determinacy | Flower GO:0010093 | specification of floral organ identity | 8 |
| GO:0010227 | floral organ abscission | Flower GO:0048833 | specification of floral organ number | 2 |
| GO:0048437 | floral organ development | Flower GO:0048437 | | 92 |
Table 2 showed MADS-box and YABBY transcription factors identified in different wheat flower substructures. MElightskeblue2, MElightblue4 and MEcoral1 highly correlated to wheat booting spike, had not identified any genes of MADS-box and YABBY transcription factors. MEdeeppink2 module had none of MADS-box and YABBY transcription factors, which was high correlation module of ear spike. The Rachis correlated module of MEdarkgoldenrod1 also had not identified any of MADS-box and YABBY transcription factors. However, MEchocolate4 module of early ovary identified 7 genes of YABBY family, and 6 genes of MADS-box family. This indicated that MADS-box and YABBY transcription factors were participated to regulated the development of ovary. MEdarkseagreen1 module of pistillody-stamen identified 10 of YABBY transcription factors of family, and 19 of MADS-box transcription factors. Genes high correlated to the development of stamen were grouped into MEmediumpurple3 and MEturquoise module. There were 3 transcription factors of MADS-box family identified in MEmediumpurple3 module. MEturquoise module identified 19 MADS-box transcription factors. A total of 47 MADS-box and 17 YABBY genes were found out from early ovary, pistillody-stamen, and stamen (Table 2).
Cytoscape V3.3 was utilized to visualize gene networks. Figure 3 depicted the genes networks related to Table 2.
to wheat flower development in MEdarkseagreen1 module. In Fig. 3, each node represents a gene. The lines between nodes represent the correlations. The networks showed the genes networks of wheat anther (Fig. 3A), style (Fig. 3B), filament (Fig. 3C), and ovary (Fig. 3D), respectively. In Fig. 3, the larger the node, the more connectivity its input gene had in the networks. This indicated that they had more important roles involved in wheat flower development. There were twelve hub genes in blue color associated anther development, such as TraesCS4A02G059000.1, TraesCS3D02G425800.1, TraesCS4D02G235800.1 (Fig. 3A). Genes of TraesCS4A02G058800.3, TraesCS4A02G058800.1, and TraesCS4A02G058800.5 in pink color had greater connectivity than other genes, demonstrating that they had more regulating role with style development (Fig. 3B). Figure 3C displayed that genes of TraesCS5B02G202000.1, TraesCS5A02G203300.1, and TraesCS5D02G209700.1 played a significant role in filament development. There were 10 hub genes affected ovary development showing in Fig. 3D, such as TraesCS2A02G514200.1, TraesCS2B02G542400.2, TraesCS4A02G422700.1.

To validate genes regulated wheat pistillody-stamen development, the expression profiles of twenty genes were examined by qRT-PCR (Fig. 4). In Fig. 4, the expression levels in pistillody-stamen of the 20 genes were apparently greater than in stamen. This suggested that these data identified by WGCNA from RNA-seq were reliable, and these genes played important roles in wheat pistillody-stamen development. The expression level of the ten genes in wheat pistillody-stamen was more than ten times greater than that in stamen, which ID were as follows: TraesCS6B02G411300.1, TraesCS5D02G294500.1, TraesCS4D02G130400.3, TraesCS4A02G058800.3, TraesCS4A02G058800.1, TraesCS4A02G058800.5, TraesCS4B02G245900.5, TraesCS4B02G245900.4, TraesCS5A02G155100.1, TraesCS5A02G286800.1. The following ten genes had very low expression levels in stamen than in pistillody-stamen, indicating that they might not be associated with stamen development: TraesCS6B02G411300.1, TraesCS5D02G294500.1, TraesCS2A02G514200.1, TraesCS4D02G130400.3, TraesCS6B02G012800.1, TraesCS4A02G058800.1, TraesCS4A02G058800.5, TraesCS4B02G245900.5, TraesCS4B02G245900.4, and TraesCS5A02G155100.1.

Discussion

WGCNA can identify hub genes from tissue-specific modules, explore gene interaction networks, and
carry out functional annotation about unknown genes, according to genes with similar functions always have similar expression levels [21]. Therefore it has been widely used to reveal genes interaction networks of flower development in different plants [22, 23, 25]. Particularly, Ramírez-González et al. established tissue- and stress-specific co-expression networks of wheat Azhurnaya cultivar, including spikelet development networks. These networks and gene expression atlases were shown in wheat eFP browser (http://bar.utoronto.ca) [1]. However, the identification of wheat pistillody-stamen related genes and their interaction networks by WGCNA has never been reported. In this research, WGCNA was applied to analyze the RNA-sequence data of 11 wheat tissues, including pistil, pistillody-stamen of HTS-1, and stamen of CSTP. The result of WGCNA Fig.d out that the MEdarkseagreen1 module significantly associated with the development of pistillody-stamen of HTS-1 (Fig. 2). Then a total of 180 genes correlated to wheat flower development were identified in this module by GO analysis (Additional file 1: Table S1). We have queried for the information of 180 wheat flower development associated genes in wheat eFP browser. The query results in wheat eFP browser has proved that all of the 180 genes are expressed in wheat spikelet, which suggest that they are affected with the development of wheat flower. Hence the query results prove that WGCNA is an efficient and accurate analysis method to cluster genes from high throughput data. However networks in wheat eFP browser don’t reveal and provide information about mechanism of genes for pistillody-stamen development. Because Azhurnaya cultivar is not a pistillody-stamen wheat. WGCNA results of this research can not only reveal the gene modules associated pistillody-stamen development, but also identify specific gene modules involving in other wheat tissues, except pistil and late ovary (Fig. 2). The MEdarkseagreen1 module is highly related to wheat pistillody-stamen (Fig. 2). Moreover, gene number related HTS-1 style and ovary development is 5 and 40, respectively (Additional file 1: Table S1). In order to verify the accurate of candidate genes related to pistillody-stamen development, we selected 20 candidate genes to verify their expression pattern by qRT-PCR (Fig. 4). In Fig. 4, 14 proteins of 20 sequences are related to ovary development of pistillody-stamen, and 5 of 20 proteins are associated with the development of style in pistillody-stamen (Additional file 1: Table S1). Among 14 ovary genes mentioned above, 7 gene are hub genes showing in Fig. 3D, which IDs are
as follows: TraesCS2A02G514200.1, TraesCS6B02G012800.1, TraesCS4D02G101700.1, TraesCS7D02G060400.2, TraesCS3D02G442000.5, TraesCS2B02G542400.2, TraesCS5A02G155100.1. The pistillody and pistil expression level of the 7 hub genes of ovary development are apparently higher than that of in stamen (Fig. 4). This result infer that the 7 hub genes must participate to regulate the development of ovary. Among 5 style associated genes, 3 genes are hub gene related the development of style; Their IDs are as follows: TraesCS4A02G058800.3, TraesCS4A02G058800.1, TraesCS4A02G058800.5 (Fig. 3B). The expression level of the 3 style related hub genes are barely expressed in stamen, and their pistillody-stamen expression are greater than in stamen (Fig. 4). This indicate that TraesCS4A02G058800.3, TraesCS4A02G058800.1, TraesCS4A02G058800.5 must regulate the development of pistillody-stamen. qRT-PCR results show that the prediction genes associated with pistillody-stamen development by WGCNA are very reliable (Fig. 4).

The development of flower features directly affect crop yield and other agronomic traits. Male sterility has been widely studied in flower development, and has very important application value in breeding [8]. Stamen transformation into pistil or pistil-like structure (called pistillody-stamen) is rare phenomenon of male sterility. Previous studies have reported some pistillody-stamen material in wheat, such as (cr)-n26, (cr)-csdt7bs, and HTS-1[4, 5, 8], which let us have some insight into understanding the molecular interaction mechanism of wheat pistillody-stamen genes [5, 10–12]. We compare the published pistillody related proteins with proteins of HTS-1 by BLAST. The alignment results are deposited in Additional file 2: Table S2. The BLAST results display that protein of WAP3 gene (NCBI Accession: BAA33459.1) [5] have two highly similar proteins in HTS-1, which IDs are TraesCS7B02G286600.2 and TraesCS7A02G383800.1 (Additional file 2: Table S2). Both of the identity value and the ratio of alignment sequence cover the total length between WAP3 gene and HTS-1 genes are greater than 95% (Additional file 2: Table S2). Murai et al. reported that the expression level of WAP3 decrease could induce the pistillody of (cr)-csdt7bs [5]. TraesCS7B02G286600.2 and TraesCS7A02G383800.1 of HTS-1 are confirmed to be expressed in stamen of HTS-1 (Additional file 2: Table S2). Therefore this result is conform with the published results of Murai et al.. This suggest that TraesCS7B02G286600.2 and TraesCS7A02G383800.1 down regulate the development of pistil, and
will maintain a certain expression level in stamen. When its amount of expression decrease below a certain level, the stamen will transform into pistil. The mitochondrial gene orf25 protein (BAA82046.1) [9] are highly similar with TraesCS4D02G236200.1 and TraesCS7A02G099300.1 of HTS-1. This result suggest that TraesCS4D02G236200.1 and TraesCS7A02G099300.1 can induce pistillody of HTS-1.

Protein of orf260^{cra} gene (BAG84632.1) is associated with pistillody induction [10]. BAG84632.1 has two similarity sequences in HTS-1 (TraesCS4D02G140600.1 and TraesCS7A02G237800.1) (Additional file 2: Table S2). The identity between orf260^{cra} and two HTS-1 sequences mentioned above are more than 95% (Additional file 2: Table S2). This result imply that TraesCS4D02G140600.1 and TraesCS7A02G237800.1 might be associated the development of pistillody-stamen development. TraesCS4D02G140600.1 and TraesCS7A02G237800.1 will be used as candidate gene for further study. Protein of WCNP1 gene (BAM65843.1) was significantly up-regulated in young spikes of the pistillody-stamens in (cr)-csdt7bs induced by mitochondrial retrograde signaling [11]. BAM65843.1 shows high similarity with HTS-1 proteins of TraesCS5D02G102100.1 and TraesCS5B02G095800.1, whose identity and ratio are greater than 95% (Additional file 2: Table S2). Whereas the function of TraesCS5D02G102100.1 and TraesCS5B02G095800.1 are still unclear in HTS-1. Thus TraesCS5D02G102100.1 and TraesCS5B02G095800.1 will be used as candidate genes associated the development of pistillody-stamen development for further study. In 2015, Yang et al. reported 167 proteins of 206 DEG proteins related to flower development of HTS-1 [12]. The 167 protein of 206 DEG have been aligned against protein of HTS-1 identified by WGCNA. 123 of 167 protein sequences are highly similar with 214 proteins of HTS-1 (Additional file 2: Table S2). The identities and ratios between 123 DEG and 214 records are above 95% and 90%, respectively (Additional file 2: Table S2). 43 proteins of the 123 proteins are expressed in stamen of HTS-1, and 50 proteins of 123 proteins are found in pistillody-stamen of HTS-1 (Additional file 2: Table S2). This results display that 50 proteins expressed in pistillody-stamen of HTS-1 might be associated with the development of pistillody-stamen in HTS-1. In 2019, Sun et al. have reported that TaEPFL1 gene(comp109492_c0) plays an important role in the development of stamen and that over-expression of TaEPFL1 results in pistillody-
stamens [13]. The alignment identity of TaEPFL1 gene against TraesCS6D02G296500.1 and
TraesCS6B02G347500.1 were 100%. This result is in accord with the result of Sun et al. [13]. Finally, a
total of 58 pistillody-stamen development associated proteins are validated by BLAST in HTS-1. These
58 sequences will be considered as pistillody-stamen development candidate genes for future studies.
Transcription factors play an important role in regulating the development of plants. Some previous
studies have reported that MADS-box [9, 14, 16, 17, 26, 27, 28] and DROOPING LEAF gene of YABBY
family [6, 15, 16] associated with the development of wheat pistillody-stamen. We identify MADS-box
and YABBY transcription factors from wheat flower substructure tissues (Table 2). Totally, there are 47
MADS-box identified in early ovar, pistillody-stamen, and stamen (Table 2). MADS-box transcription
factors identified in early ovary, pistillody-stamen, and stamen are different. Similarly, YABBY
transcription factors of early ovary are different from those in pistillody-stamen. This indicate that
different MADS-box and YABBY genes regulating the development of different flower sub-structure,
according to ABCED model of flower development [6]. According to the information of 47 MADS-box,
we further identify 10 MADS-box transcription factors from the 180 proteins in Tables S1, which IDs
are showed as follows: TraesCS1A02G199600.1, TraesCS5A02G286800.1, TraesCS5D02G294500.1,
TraesCS6A02G158100.1, TraesCS6A02G259000.1, TraesCS6A02G259000.3, TraesCS6B02G186700.1,
TraesCS6B02G286400.2, TraesCS6D02G147400.1, TraesCS6D02G240200.2. In Fig. 4, the results of
qRT-PCR show that the expression levels of TraesCS5A02G286800.1 and TraesCS5D02G294500.1 are
significantly higher in pistillody-stamen and pistil of HTS-1 than in stamen of CSTP. This indicates that
TraesCS5A02G286800.1 and TraesCS5D02G294500.1 participate in the development of wheat
pistillody-stamen. For identifying the unpublished genes from 47 MADS-box genes, 93 published
MADS-box genes relating to pistillody-stamen development were aligned against 47 MADS-box genes
by BLAST. Additional file 3: Table S3 showed that 11 of 47 MADS-box gene had exactly the same as
the published sequences. This result implied that the MADS-box genes of wheat were highly
conserved. Therefore, the remaining 36 of 47 MADS-box gene have not been reported yet.
A total of 4 published YABBY transcription factors are considered to regulate the development of
wheat pistillosy-stamen [6, 15, 16]. The published 4 proteins of YABBY have been aligned with the
180 floral proteins of Additional file 1: Table S1 by BLAST. The alignment results are shown in Additional file 3: Table S3. In Additional file 3: Table S3, DROOPING LEAF gene of TaDL3 sequence is complete the same as proteins of TraesCS4B02G245900.4 and TraesCS4B02G245900.5 of HTS-1. Protein of BAH83538.1 [6, 15] are exactly same as sequences of TraesCS4A02G058800.1 of HTS-1. Furthermore, TraesCS4A02G058800.3, TraesCS4B02G245900.4 and TraesCS4B02G245900.5 have been verified highly expressed in pistil or pistillody-stamen of HTS-1 by qRT-PCR (Fig. 4). This results suggest that TraesCS4A02G058800.3 and TraesCS4B02G245900.4 played important role in pistillody-stamen and pistil development. Thus these result show that 14 proteins of 17 of YABBY have never been reported, excluding TraesCS4A02G058800.3, TraesCS4D02G245300.3, and TraesCS4D02G245300.2 (Table 2). The 14 YABBY proteins will be used as candidate genes to study the mechanism of pistillody-stamen development.

Conclusions
WGCNA are applied to explore genes related to pistillody-stamen development and their interaction networks. In the current study, we identify 10 and 3 hub genes play an important role in filament and ovary development of HTS-1. A total of 58 pistillody-stamen development associated proteins are validated by BLAST. 36 of 47 MADS-box and 14 of 17 YABBY transcription factors are considered to regulate the development of pistillody-stamen, which have not been reported yet. Future studies of these results will provide further insights into the pistillody-stamen development of HTS-1.

Materials And Methods
Data source
We obtained the following wheat RNA-seq data from the National Center for Biotechnology Information (NCBI) Sequence Read Archive database (SRA, http://www.ncbi.nlm.nih.gov/sra): root (ERR424732), stem (ERR424762), flag leaf (SRR3068439), booting spike (SRR6802610), ear spike (SRR6802611), rachis (SRR6802608), early ovary (SRR6802613), late ovary (SRR6802612), HTS-1 pistil (SRR1175868), HTS-1 pistillody stamen (SRR1177760), and CSTP stamen (SRR1177761). Wheat assembled genomic DNA sequences, coding sequences (CDS), and protein sequences were downloaded from the Ensembl Plants database (version release 41,
ftp://ftp.ensemblgenomes.org/pub/plants/release-41/). Totally, there were 133,346 CDS and protein sequences.

Analysis of gene quantification

Salmon [29] was utilized to quantify gene expression of 11 wheat tissues mentioned above, and 133,346 CDS were used as reference sequences. We filtered out the gene quantification of transcripts per million (TPM) values less than 0.5 in pistillody stamen, and there were 58,576 genes remained. A total of 58,576 genes with TPM values in 11 wheat tissues were obtained.

WGCNA analysis

The WGCNA R package [30] was applied to construct the co-expression network of 11 wheat tissues (58,576 genes). Firstly, genes with mean TPM value less than 1 were filter out, and there were 47,492 genes remained. In order to make the genes in the network conform to the scale-free topology distribution, a soft threshold (power) must be set. When scale-free topology fitted index of 0.9, the first soft threshold (power) of 18 was used (Fig. 1A). Next the adjacency matrix was converted into topological overlap matrix (TOM), and the function dissTom = 1–TOM was used to invert the TOM into their invert matrix. Then Dynamic Tree Cut was adopted for module identification, and the minimum module size was 30. The function hclust was utilized to merge the similar modules (mergeCutHeight = 0.25) (Fig. 1B). To identify the tissue specific modules, we analyzed the correlation between the tissue and module eigengene (ME) (Fig. 2). The function corPvalueStudent() was employed to do the significance test between the trait and module eigengene.

GO analysis and visualization

The result of WGCNA identified that the MEdarkseagreen1 module was significantly related to the development of pistillody-stamen of HTS-1. Gene Ontology (GO) analysis was planned to find out genes related to pistillody-stamen development in MEdarkseagreen1 module. A total of 4,702 protein sequences were obtained from the Ensembl Plants database (version release-41), which edge values were more than 0.3 in MEdarkseagreen1 module. Kobas 3.0 was applied to do GO annotation of the 4,702 proteins [31]. Additional file 1: Table S1exhibited genes related wheat pistillody-stamen development by Kobas 3.0. According to GO analysis results showed in Additional file 1: Table S1,
genes involved in wheat anther, stigmas, style, filament, and ovary development were collected from MEdarkseagreen1 module. Table 1 displayed the genes number of different GO terms associated with wheat flower development in pistillody-stamen, which were accounting from Additional file 1: Table S1. Then, for identifying the unpublished proteins about pistillody-stamen development of HTS-1, we aligned the published pistillody related proteins against HTS-1 proteins by BLAST. The alignment results were deposited in Additional file 2: Table S2.

Cytoscape V3.3 [32] was used to generate genes network visualization of different flower tissues (Fig. 3). Genes with higher connectivity were considered as hub genes [1].

Transcription factors analysis

To seek for the transcription factors about pistillody-stamen development, we identified MADS-box and YABBY transcription factors from specific modules highly correlated to different flower tissues using the method of Chen et al. [33]. The MADS-box and YABBY transcription factors for different flower tissues were shown in Table 2. In order to find out unpublished transcription factors related to pistillody-stamen development, we utilized BLAST program to compare the published wheat MADS-box and YABBY transcription factors related to pistillody with HTS-1 transcription factors in Table 2. The BLAST results were displayed in Additional file 3: Table S3.

Plant materials and RNA extraction

Both CSTP and HTS-1 are common wheat maintained in our laboratory. CSTP is a near-isogenic line of Chinese Spring, which has three pistils in each floret [8]. HTS-1 is a pistillody wheat mutant derived from CSTP [8]. CSTP and its pistillody mutant of HTS-1 were planted in a field in the China West Normal University, Nanchong, China. Pistils and pistillody-stamens of HTS-1 were collected at the heading stage, and stamens of CSTP were harvested at the same time. Each samples has two biological replicates. Total RNA of the pistils and pistillody-stamens of HTS-1, and stamens of CSTP were extracted using a Plant Easy Spin RNA Miniprep Kit (Biomiga, China). DNase treatment was performed before proceeding with complementary DNA (cDNA) synthesis. The quality of the cDNA was determined by gel electrophoresis. cDNA concentration was calculated and adjusted to 50 ng/µl by spectrophotometer (Thermo, Nanodrop 2000).
Quantitative real-time PCR analysis
The qRT-PCR assays were performed with SsoFast EvaGreen Supermix (Bio-Rad, USA) using Bio-Rad CFX96 real-time PCR platform (Bio-Rad, USA). The primers for the 20 genes associated with wheat flower development were shown in Additional file 4: Table S4. Each reaction contained 5 µl of SSofast (Bio-Rad), 1 µl of gene-specific primers, 0.5 µl of cDNA, 3.5 µl of ddH₂O, and in a final volume of 10 µl.

The wheat genes (GenBank No. AB181911) and wheat housekeeping genes Ubiq (DQ086482) were used as a reference gene [34]. Fold-changes of RNA transcripts were calculated via the 2-ΔΔCt method [35]. Expression pattern of 20 genes associated pistillody-stamen development were displayed in Fig. 4.

Abbreviations
CDS
Coding sequences; DEG:differentially expressed genes; GO:Gene Ontology; qRT-PCR:Quantitative real-time PCR; SRA:Sequence Read Archive; TP:three pistil; TPM:Transcripts per million;
WGCNA:Weighted gene co-expression network analysis

Declarations
Authors’ contributions
MLL participated in the design of this work, carried out the qRT-PCR experiments, and wrote the manuscript. ZYC and ZSP planned the study and designed this work, and wrote the manuscript. ZJY participated in data analysis and wrote the manuscript. WYC and SHW participated in the design of this work and contributed to writing the manuscript. All authors read and approved the final manuscript.

Funding
This study was supported by National Natural Science Foundation of China (Grant No. 31760425), the Innovation Team Project of China West Normal University (Grant No. CXTD2018-6), the Doctoral Scientific Research Program of China West Normal University (16E011), and the Meritocracy Research Funds of China West Normal University (17YC321 and 17YC322). The funding agencies were not involved in study design, data collection and analysis, or preparation of the manuscript.

Availability of data and materials
The RNA-seq datasets can be obtained from the NCBI SRA database, which have been mentioned in
materials and methods of this study. All datasets generated during this study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interest.

Acknowledgements
Not applicable.

References
1. Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, et al. The transcriptional landscape of polyploid wheat. Science. 2018;361:eaar6089.

2. United States Department of Agriculture: World Agricultural Supply and Demand Estimates. Report No. WASDE-0719. 2019. https://downloads.usda.library.cornell.edu/usda-esmis/files/3t945q76s/ks65hp42q/vd66w940f/latest.pdf.

3. The International Wheat Genome Sequencing Consortium (IWGSC). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science. 2014;345:1251788.

4. Murai K, Tsunewaki K. Photoperiod-sensitive cytoplasmic male sterility in wheat with Aegilops crassa cytoplasm. Euphytica. 1993;67:41–8.

5. Murai K, Takumi S, Koga H, Ogihara Y. Pistillody. homeotic transformation of stamens into pistil-like structures, caused by nuclear-cytoplasm interaction in wheat. Plant J. 2002;29:169–82.

6. Murai K. Homeotic genes and the ABCDE model for floral organ formation in wheat. Plants. 2013;2:379–95.

7. Peng ZS. A new mutation in wheat producing three pistils in a floret. J Agro Crop Sci. 2003;189:270-2.

8. Peng ZS, Yang ZJ, Ouyang ZM, Yang H. Characterization of a novel pistillody mutant in common wheat. Austri J Crop Science. 2013;7:159–64.

9. Hama E, Takumi S, Ogihara Y, Murai K. Pistillody is caused by alterations to the class-B MADS-box gene expression pattern in alloplasmic wheats. Planta. 2004;218:712–20.

10. Zhu Y, Sarake T, Yamamoto Y, Hagita H, Takumi S, Murai K. orf260cra, a novel mitochondrial gene, is associated with the homeotic transformation of stamen into pistil-like structures (pistillody) in alloplasmic wheat. Plant Cell
11. Yamamoto M, Shitsukawa N, Yamada M, Kato K, Takumi S, Kawaura K, Ogihara Y, Murai K. Identification of a novel homolog for a calmodulin-binding protein that is upregulated in alloplasmic wheat showing pistillody. Planta. 2013;237:1001–13.

12. Yang ZJ, Peng ZS, Wei SH, Liao ML, Yu Y, Jang ZY. Pistillody mutant reveals key insights into stamen and pistil development in wheat (Triticum aestivum L.). BMC Genom. 2016;16:211.

13. Sun QX, Qu JP, Yu Y, Yang ZJ, Wei SH, Wu YL, Yang J, Peng ZS. TaEPFL1, an EPIDERMAL PATTERNING FACTOR-LIKE (EPFL) secreted peptide gene, is required for stamen development in wheat. Genetica. 2019;147:121–30.

14. Meguro A, Takumi S, Ogihara Y, Murai K. WAG, a wheat AGAMOUS homolog, is associated with development of pistil-like stamens in alloplasmic wheats. Sex Plant Reprod. 2003;15:221–30.

15. Ishikawa M, Ohmori Y, Tanaka W, Hirabayashi C, Murai K, Ogihara Y, Yamaguchi T, Hirano HY. The spatial expression pattern of DROOPING LEAF orthologs suggest a conserved function in grasses. Genes GenetSyst. 2009;84:137–46.

16. Zhang L, Yang ZJ, Peng ZS, Yu Y, Sun QX. Molecular cloning, characterization, and expression analysis of DROOPING LEAF gene in the pistillody line of common Wheat. Modern Appl Sci. 2015;9:44–52.

17. Wang QH, Yang ZJ, Wei SH, Jiang ZY, Yang YF, Hu ZS, Sun QX, Peng ZS. Molecular cloning, characterization and expression analysis of WAG-1 in the pistillody line of common wheat. Genet Mol Res. 2015;14:12455–65.

18. Zhu XX, Li QY, Shen CC, Duan ZB, Yu DY, Niu JS, Ni YJ, Jiang YM. Transcriptome analysis for abnormal spike development of the wheat mutant dms. PLoS ONE. 2015;11:e0149287.

19. Greenside P, Shimko T, Fordyce P, Kundaje A. Discovering epistatic feature interactions from neural network models of regulatory DNA sequences. Bioinformatics. 2018;34:629–37.

20. Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, et al. A Human protein-protein interaction network: a resource for annotating the proteome. Cell. 2005;122:957–68.

21. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol. 2005;4:17.

22. Hollender CA, Kang CY, Darwish O, Geretz A, Matthews BF, Slovin J, Alkharouf N, Liu ZC. Floral transcriptomes in woodland strawberry uncover developing receptacle and anther gene networks. Plant Physi. 2014;165:1062–75.

23. Roberts WR, Roalson EH. Comparative transcriptome analyses of flower development in four species of Achimenes (Gesneriaceae). BMC Genom. 2017;18:240.

24. Ma SS, Ding ZH, Li PH. Maize network analysis revealed gene modules involved in development, nutrients utilization, metabolism, and stress response. BMC Plant Biol. 2017;17:131.

25. Shahan R, Zawora C, Wight H, Sittmann J, Wang WP, Mount SM, Liu ZC. Consensus coexpression network analysis
identifies key regulators of flower and fruit development in wild strawberry. Plant Physiol. 2018;178:202-16.

26. Yang W, Lou X, Li J, Pu M, Mirbahar AA, Liu D, Sun J, Zhan K, He L, Zhang A. Cloning and functional analysis of MADS-box genes, TaAG-A and TaAG-B, from a wheat K-type cytoplasmic male sterile line. Front. Plant Sci. 2017: 1081.

27. Su YL, Liu JX, Liang WQ, Dou YH, Fu RF, et al. Wheat AGAMOUS LIKE 6 transcription factors function in stamen development by regulating the expression of Ta APETALA3. Development. 2019;146:dev177527.

28. Paolacci AR, Tanzarella OA, Porceddu E, Varotto S, Ciaffi M. Molecular and phylogenetic analysis of MADS-box genes of MIKC type and chromosome location of SEP-like genes in wheat (Triticum aestivum L.). Mol Genet Genomics. 2007;278:689-708.

29. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. Nature Meth. 2014;14:417-9.

30. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

31. Xie C, Mao XZ, Huang JJ, Ding Y, Wu JM, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 2011;39:W316–22.

32. Shannon P, Markie A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498-504.

33. Chen ZY, Guo XJ, Chen ZX, Chen WY, Liu DC, Zheng YL, Liu YX, Wei YM, Wang JR. Genome-wide characterization of developmental stage- and tissue-specific transcription factors in wheat. BMC Genom. 2015;16:125.

34. Liao ML, Peng ZS, Yang ZJ, Wei SH, Martinek P. Identification of differentially expressed genes in a pistillody common wheat mutant using an annealing control primer system. Genet Mol Res. 2015;14:3995–4004.

35. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCt method. Methods. 2001;25:402-8.

Figures
Co-expression network analysis of 11 RNA-seq datasets for wheat flower development. A: Determination of soft threshold of WGCNA; B: Modules identified by the dynamic tree cut and the merged modules with similar expression pattern.
Module-trait relationships

Figure 2

Heatmap of module-trait relationships. Each row in the heatmap represents a module. Each column represents a wheat tissue. The color of each cell represents the association coefficient between the module and the tissue type. Red color of each cell indicates the positive relationships between module and tissue type. Green color of each cell implies the negative relationships between module and tissue type.
Co-expression network visualization of pistillody-stamen development genes in MEDarkseagreen1 module. The networks depict gene relationships of wheat anther (A), style (B), filament (C), and ovary (D) development. Each node represents a gene, and lines between genes represent co-expression relationships. The size of nodes imply the connectivity of the input genes.
Figure 4

Expression patterns of twenty pistillody-stamen development genes verified by real-time PCR. P represents fully developed wheat pistill of HTS-1; PS represents completely developed wheat pistillody-stamen of HTS-1; S represents fully developed wheat stamen of CSTP.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

Table S3.xlsx
Table S4.xlsx
Table S1.xlsx
Table S2.xlsx