**Methods**

*RNA-seq datasets*

Raw sequence reads for 48 cotton seed RNA-seq datasets were obtained from the NCBI BioProject PRJNA179447, including those from two accessions of allopolyploid cotton *Gossypium hirsutum* (elite cultivar TM1 and wild race Yuc), and its model diploid progenitors - *G. arboreum* (A2) and *G. raimondii* (D5). For each seed sampling condition, equal amount of diploid A2 and D5 reads were combined to generate *in silico* allopolyploid datasets (ADs). As previously described in Hovav *et al.* (2015) and Hu *et al.* (2016), raw reads were filtered using the FASTX-Toolkit ([http://hannonlab.cshl.edu/fastxtoolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)) and mapped to the reference diploid D5 genome (Paterson et al. 2012) using GSNAP (Wu and Nacu 2010) with single-nucleotide polymorphism (SNP) tolerance mapping between the A2 and D5 diploids or their co-resident counterparts in allopolyploids (Page and Udall 2015). Following mapping, reads containing SNPs that are diagnostic of sub-genome origins were partitioned into homoeologous At and Dt counts by PolyCat (Page et al. 2013). For comparison, HyLite (Duchemin et al. 2015) was used as a secondary pipeline for estimating homoeologous expression levels from filtered RNA-seq reads in a single step. Briefly, the HyLiTE pipeline recruited Bowtie2 (Langmead and Salzberg 2012) to map reads to a reference diploid genome and automatically identified SNPs diagnostic of parental lineages, based upon which polyploid reads were classified to parental origins and summarized into homoeolog-specific read count tables. A third pipeline was built upon default Bowtie2 (Langmead and Salzberg 2012) mapping against a transcriptomic reference containing both homoeologous gene models, followed by RSEM (Li and Dewey 2011) quantification. The transcriptomic reference was constructed from the reference diploid D5 genome and the homoeologous SNP index as used in the PolyCat pipeline.

*Performance assessment of homoeolog expression estimation*

Three measures were defined to evaluate the performance of homoeolog-specific read assignment – *Efficiency* (***Ef***), *Accuracy* (***Ac***), and *Discrepancy* (***Di***). As reads mapped to a diploid reference gene model represent the total expression (T) for a pair of homoeologous genes (At and Dt, t denotes sub-genome origin), these reads include those diagnostic of homoeologs and reads indistinguishable between subgnomes (N); thus, T = At + Dt + N. *Efficiency* measures the percentage of total reads can be assigned to homoeologs as ***Ef* =** . According to the known sub-genome origin of the *in silico* synthetic allopolyploid (ADs) reads, correctly assigned homoeolog-specific reads can be recovered from At and Dt (At = At.true + Dt.false; Dt = Dt.true + At.false). Assignment *Accuracy* is defined as ***Ac* =** . Considering that inaccurate assignment of both At and Dt reads may cancel out each other and results in no difference between observed and expected read counts, we defined *Discrepancy* as a measure of absolute deviation from expected counts (At.exp and Dt.exp): ***D*** = . In addition to measures aggregated for both homoeologs, separate assessments for At and Dt were performed accordingly (Supplemental Table ?). Use Correlations among variables and Multiple Logistic Regression tests to show statistical evidences for the association between performance metrics and explanatory variables. R scripts for performance assessment and below analyses are available at https://github.com/???.

**Supplemental table ?: Definition of performance measures of homoeolog expression estimation.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Aggregated | *At only* | *Dt only* |
| *Efficiency* (*Ef*) |  |  |  |
| *Accuracy* (*Ac*) |  |  |  |
| *Discrepancy* (*Di*) |  |  |  |

*Gene expression analysis*

Two ~~Three~~ methods were applied to analyze differential expression (DE) of homoeologs for each polyploid sample condition. Both using a negative binomial model to estimate gene-wise dispersion parameters, DESeq2 (Love et al. 2014) takes a classical hypothesis testing approach to report nomial p-values, while EBSeq (Leng et al. 2013) is cast within a Bayesian framework and reports posterior probabilities for each gene to be equally or differentially expressed. To identify differential expression at a false discovery rate α=0.05, the DESeq2 p-values were adjusted for multiplicities using Benjamini–Hochberg method ([Benjamini and Hochberg, 1995](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3554454/#bib3)), and the posterior probability of DE by EBSeq was considered >1- α. ~~A third method was applied using fisher’s exact tests as previously described (Yoo et al. 2013). Briefly, replicates were normalized by total library size, and tests were restricted to genes having read counts in all replicates in each contrast; the distribution of P-values was controlled for a false discovery rate by the BH method at α=0.05.~~ To evaluate the discriminative performance of a DE method for polyploid datasets, DE genes inferred between were parental diploids A2 and D5 were used as true labels, to categorize true positive and false positive DE results from synthetic allopolyploid (ADs). Sensitivity (true-positive rate) and specificity (true-negative rate) were calculated for each method.

*Gene co-expression analysis*

For data transformation, RPKM followed by log2 transformation were performed with local R script, and regularized logarithm *rlog* transformation were conducted with R package DESeq2 (Love et al. 2014). Based on transformed expression profiles, Pearson correlation coefficient was calculated for each pair of genes, and Fisher’s z-test was used to identify significant differential correlations (DC) with *P* < 0.05. DC gene pairs were classified as having gain of correlation (GOC) or loss of correlation (LOC) against the reference condition. Briefly, based on a threshold of *P* < 0.05 for correlation significance and the sign of correlation in each condition, gene-gene correlations in each condition can be categorized into 3 classes, i.e. significant positive correlation (+), no significant correlation (0), and significant negative correlation (-); thus, a total of 9 classes can be inferred between two conditions. For example, class 0/+ indicates GOC from a not significant correlation (0) to a significant positive correlation (+), and class +/- indicates LOC from a positive correlation to a negative correlation. Identification and classification of DC genes pairs were performed using the R package DGCA (McKenzie et al. 2016). To determine if a specific DC class is significantly enriched, one-sided Fisher's exact tests was applied to evaluate the overlap between significant (*P* < 0.05) and all gene pairs classified into this class.

Differentially co-expressed genes were identified as those significantly over-represented by DC gene pairs. Briefly, the percentage of DC among all possible gene pairs measures the extent of differential co-expression, or the probability *p* to detect “a differential co-expression gene pair”. For a gene observed in *k* differential correlation pairs among all possible pairs *n*, the probability *P* of a “differential co-expression gene” follows the binomial distribution model as follows:

*P* was corrected by the BH method (Benjamini and Hochberg 1995) at α = 0.05 for identifying differentially co-expression genes.

*Co-expression network construction*

The construction of a co-expression network begins with a measure of similarity score from pairwise expression profiles for each pair of genes, and the next step is converting the similarity scores to an adjacency matrix that reflects connections between gene nodes in network (Serin et al. 2016). By applying a threshold to similarity scores to determine whether a pair of gene should be connected in the network, an adjacency matrix containing only 0 and 1 values is resulted as a binary network. Instead of hard thresholds, weighted gene co-expression network analysis (WGCNA) applies a soft threshold chosen by the scale-free topology criterion to assign a connection weight to each gene pair, which allows the adjacency matrix to present network connections quantitatively (Zhang and Horvath 2005; Langfelder and Horvath 2008). Using Pearson correlation coefficient to determine similarity scores, we constructed a series of binary and weighted network for each expression dataset generated from three read estimating pipelines followed by log2 RPKM or *rlog* data transformation. Two types of hard threshold were used to generate binary networks – rank based and Fisher’s Z statistics based thresholds. Based on the ranking of similarity scores, a set of cutoffs (5%, 1%, 0.5% and 0.1%) were applied to select top ranked connections as edges in network. Fisher’s Z transformation normalizes the distribution of Pearson correlation and converts each correlation coefficient to a z-score for significance test. A set of Z cutoffs (1.5, 2.0, 2.5, 3.0) were applied to select significant correlations as edges in network. To construct weighted networks, we also used a test of soft thresholds (0, 12, 24) to test the effect of thresholding.

Binary network construction is a decision problem as the inference algorithm decides whether an edge exists (positive) between a pair of genes or not (negative). Treating the edges inferred from true homoeolog expression data as true label, we can evaluate the performance of binary network construction to recover true edges from estimated homoeolog expression data with a Receiver Operating Characteristic (ROC) curve. Due to the large gene number of polyploid networks (> 60,000 genes), a 10% random sampling of genes was used to compute ROC curves and AUC (area under receiver characteristic curve) measures with 10 repeats. AUC of 0.5 indicates a completely random guess, 0.7 is generally good and 0.9 or above is very high.

*Network topology and functional connectivity assessment*

Node connectivity (*k*) measures the amount of connections to a given node in network, which is calculated as the summation of connected edge numbers or edge weights. In terms of functional connectivity of the network, the ‘Guilt-by-Association’ principle suggests that genes of similar functions are preferentially connected in network. Therefore, a neighbor voting algorithm can be used to classify genes as belonging to a given functional group according to the functionality of their connected genes (i.e. neighborhood). Based upon the known functional labels of genes (here we used GO and KEGG annotations), the voting algorithm acts as a binary classifier to return TRUE or FALSE predictions of functional labels, whose performance can be assessed with a Receiver Operating Characteristic (ROC) curve. The derived metric of AUC (area under receiver characteristic curve) characterizes the degree to which an input network topology can predict the gene membership of a functional category, which intuitively corresponds to the assessment of functional connectivity. GO and KEGG terms were extracted from the v2.1 annotation of *Gossypium raimondii* reference genome downloaded from Phytozome ([www.phytozome.net/)](http://www.phytozome.net/)). Functional connectivity assessment was performed with R package EGAD (Ballouz et al. 2017).

Ballouz S, Weber M, Pavlidis P, Gillis J (2017) EGAD: ultra-fast functional analysis of gene networks. Bioinformatics 33 (4):612-614. doi:10.1093/bioinformatics/btw695

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological):289-300

Duchemin W, Dupont PY, Campbell MA, Ganley AR, Cox MP (2015) HyLiTE: accurate and flexible analysis of gene expression in hybrid and allopolyploid species. BMC bioinformatics 16:8. doi:10.1186/s12859-014-0433-8

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics 9:559. doi:10.1186/1471-2105-9-559

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature methods 9 (4):357-359. doi:10.1038/nmeth.1923

Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BM, Haag JD, Gould MN, Stewart RM, Kendziorski C (2013) EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics 29 (8):1035-1043. doi:10.1093/bioinformatics/btt087

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics 12:323. doi:10.1186/1471-2105-12-323

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15 (12):550. doi:10.1186/s13059-014-0550-8

McKenzie AT, Katsyv I, Song WM, Wang M, Zhang B (2016) DGCA: A comprehensive R package for differential gene correlation analysis. BMC systems biology 10 (1):106. doi:10.1186/s12918-016-0349-1

Page JT, Gingle AR, Udall JA (2013) PolyCat: a resource for genome categorization of sequencing reads from allopolyploid organisms. G3 3 (3):517-525. doi:10.1534/g3.112.005298

Page JT, Udall JA (2015) Methods for mapping and categorization of DNA sequence reads from allopolyploid organisms. BMC Genet 16 Suppl 2:S4. doi:10.1186/1471-2156-16-S2-S4

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J, Yoo MJ, Byers R, Chen W, Doron-Faigenboim A, Duke MV, Gong L, Grimwood J, Grover C, Grupp K, Hu G, Lee TH, Li J, Lin L, Liu T, Marler BS, Page JT, Roberts AW, Romanel E, Sanders WS, Szadkowski E, Tan X, Tang H, Xu C, Wang J, Wang Z, Zhang D, Zhang L, Ashrafi H, Bedon F, Bowers JE, Brubaker CL, Chee PW, Das S, Gingle AR, Haigler CH, Harker D, Hoffmann LV, Hovav R, Jones DC, Lemke C, Mansoor S, ur Rahman M, Rainville LN, Rambani A, Reddy UK, Rong JK, Saranga Y, Scheffler BE, Scheffler JA, Stelly DM, Triplett BA, Van Deynze A, Vaslin MF, Waghmare VN, Walford SA, Wright RJ, Zaki EA, Zhang T, Dennis ES, Mayer KF, Peterson DG, Rokhsar DS, Wang X, Schmutz J (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature 492 (7429):423-427. doi:10.1038/nature11798

Serin EA, Nijveen H, Hilhorst HW, Ligterink W (2016) Learning from Co-expression Networks: Possibilities and Challenges. Front Plant Sci 7:444. doi:10.3389/fpls.2016.00444

Wu TD, Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26 (7):873-881. doi:10.1093/bioinformatics/btq057

Yoo MJ, Szadkowski E, Wendel JF (2013) Homoeolog expression bias and expression level dominance in allopolyploid cotton. Heredity 110 (2):171-180. doi:10.1038/hdy.2012.94

Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Statistical applications in genetics and molecular biology 4:Article17. doi:10.2202/1544-6115.1128