Article type : Original Article

Systems genetics reveals a transcriptional network associated with susceptibility in the maize-gray leaf spot pathosystem

Nanette Christie[1,2], Alexander A. Myburg[1], Fourie Joubert[2], Shane L. Murray[3,4], Maryke Carstens[5], Yao-Cheng Lin[6,7], Jacqueline Meyer[3,5], Bridget G. Crampton[5], Shawn A Christensen[8], Jean F. Ntuli[4], Sara S. Wighard[4], Yves Van de Peer[6,7,9], Dave K. Berger[5*]

Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), nomics Research Institute (GRI), University of Pretoria, Pretoria, South Africa

Centre for Bioinformatics and Computational Biology, Genomics Research Institute, Department of Biochemistry, University of Pretoria, Pretoria, South Africa

Centre for Proteomic and Genomic Research, Cape Town, South Africa

Department of Molecular and Cell Biology, University of Cape Town, Rondebosch, South Africa

Department of Plant and Soil Sciences, Forestry and Agricultural Biotechnology Institute (FABI), Genomics Research Institute (GRI), University of Pretoria, Pretoria, South Africa

Department of Plant Systems Biology, VIB, Ghent B-9052, Belgium

[7]Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent B-9052, Belgium

Center for Medical, Agricultural, and Veterinary Entomology, United States Department of Agriculture, Agricultural Research Service, Chemistry Research Unit, Gainesville, Florida 32608

Department of Genetics, Genomics Research Institute, University of Pretoria, Pretoria 0028, South Africa

* Corresponding author

E-mail: Dave.Berger@up.ac.za

Running head: Systems genetics of maize-gray leaf spot disease

Keywords: *Cercospora zeina*, *Zea mays*, gray leaf spot, eQTL, co-expression, disease resistance, QTL, disease susceptibility, *Cercospora*

SUMMARY

We used a systems genetics approach to elucidate molecular mechanisms of maize responses to gray leaf spot (GLS) disease, caused by *Cercospora zeina,* a threat to maize production globally. Expression analysis of earleaf samples in a sub-tropical maize RIL population (CML444 X SC Malawi) subjected in field to *C. zeina* infection allowed detection of 20,206 expression QTLs (eQTL). Four *trans*-eQTL hotspots coincided with GLS disease QTLs mapped in the same field experiment. Co-expression network analysis identified three expression modules correlated with GLS disease scores. The module (*GY-s*) most highly correlated with susceptibility (r = 0.71; 179 genes) was enriched for

the glyoxylate pathway, lipid metabolism, diterpenoid biosynthesis and responses to pathogen molecules such as chitin. The *GY-s* module was enriched for genes with *trans*-eQTLs in hotspots on chromosomes 9 and 10, which also coincided with phenotypic QTLs for GLS susceptibility. This transcriptional network has significant overlap with the GLS susceptibility response of maize line B73, and may reflect pathogen manipulation for nutrient acquisition and/or unsuccessful defense responses, such as kauralexin production by the diterpenoid biosynthesis pathway. The co-expression module that correlated best with resistance (*TQ-r*; 1498 genes) was enriched for genes with *trans*-eQTLs in hotspots coinciding with GLS resistance QTLs on chromosome 9. Jasmonate responses were implicated in resistance to GLS through co-expression of *COI-1* and enrichment of genes with the GO term "cullin-RING ubiquitin ligase complex" in the *TQ-r* module. Consistent with this, JAZ repressor expression was highly correlated with GLS disease severity in the *GY-s* susceptibility network.

INTRODUCTION

Disease resistance in crop plants such as maize is often associated with one or more quantitative trait loci (QTLs), a phenomenon referred to as quantitative disease resistance (QDR) (Poland *et al.*, 2009; St.Clair, 2010). QDR is characterized by disease severity scores that exhibit a continuous distribution in a segregating population and is often strongly influenced by the environment (St.Clair, 2010; Vanderplank, 1984). This is in contrast to qualitative disease resistance conferred by allelic differences in "major genes" that have a large effect on the phenotype, and individuals that can be clearly categorized as resistant or susceptible with limited environmental influence (St.Clair,

2010; Vanderplank, 1984). Examples of qualitative resistance in maize are the major loci linked to rust disease resistance (Jines *et al.*, 2007).

A review of the genetic architecture of disease resistance in maize listed 50 publications on the mapping of maize disease resistance loci (Wisser *et al.*, 2006). However, identifying the genes, and ultimately the genetic polymorphisms, underlying QTLs that confer QDR in maize is currently a major challenge (Jamann *et al.*, 2013). Therefore, as an initial strategy towards cloning a QTL, transcriptomics can be employed to explore the molecular processes leading to resistance or susceptibility conferred by a QTL (Liu *et al.*, 2016). Furthermore, transcriptome analysis of a susceptible interaction can yield insights into both targets of pathogen manipulation as well as host defence responses as shown in a study of *Colletotrichum graminicola* and a highly susceptible maize host (Vargas *et al.*, 2012).

Gray leaf spot (GLS) is a damaging foliar disease of maize that has been documented in many sub-tropical and tropical regions of the world (Ward *et al.*, 1999). Symptoms are matchstick-shaped necrotic lesions with a gray tint on the leaf surface (Meisel *et al.*, 2009). The fungus *Cercospora zeae-maydis* is the predominant causal agent in the USA and other maize growing regions (Wang *et al.*, 1998). The related fungal species *Cercospora zeina* also causes GLS, and has been recently documented in South Africa, China and Brazil (Liu and Xu, 2013; Meisel *et al.*, 2009; Neves *et al.*, 2015).

Responses of maize to GLS pathogens have been characterized as quantitative, since there is no evidence to date of *Cercospora* spp. pathovars with cognate maize resistance genes, in contrast to other pathosystems such as maize-northern corn leaf blight (Poland *et al.*, 2009; Hurni *et al.*, 2015). More than a dozen GLS resistance QTL mapping studies have been conducted with bi-parental maize mapping populations (Berger *et al.*,

2014; Wisser *et al.*, 2006). QTL mapping of a maize RIL population derived from sub-tropical maize inbreds CML444 X SC Malawi in South Africa, where only *C. zeina* is present, identified seven GLS QTLs from five field sites (Berger *et al.*, 2014). Meta-analysis of studies to date located hotspots of GLS QTLs on chromosomes one, two, four, five and seven (Berger *et al.*, 2014).

Genome wide association mapping (GWAS) in a panel of 253 diverse inbred maize lines led to the identification of single-nucleotide polymorphisms (SNPs) in a glutathione S-transferase gene in chromosome bin 7.03 that were associated with resistance to GLS (Wisser *et al.*, 2011). GWAS of 161 inbred lines was carried out in China, and SNPs in bin 3.07 and bin 9.07 were proposed as markers for molecular breeding for GLS resistance (Shi *et al.*, 2014). The maize nested association mapping (NAM) population was recently used to identify three QTLs that reduced GLS disease by more than 10% (located in bins 1.04, 2.09 and 4.05), and experiments with near isogenic lines were conducted to develop hypotheses about resistance mechanisms (Benson *et al.*, 2015). Bi-parental QTL mapping was combined with GWAS to refine the map positions of GLS QTLs on chromosomes one, six, seven and eight (Mammadov *et al.*, 2015). A GLS QTL in bin 5.03/5.04 was fine-mapped to a region with 15 candidate genes (Xu *et al.*, 2014). Despite this, no GLS resistance genes have been cloned to date, and thus there is a need for alternative research strategies. Systems genetics is an approach to explore the molecular basis of quantitative traits such as QDR through integration of genetic data with other component molecular data gathered using -omics technologies (Civelek and Lusis, 2014).

Natural variation in transcript abundance can be used in several ways in a systems genetics approach (Feltus, 2014). The first is co-expression analysis to search for groups of genes with coordinated gene expression patterns across individuals of a

population or progeny from a cross (Zhang and Horvath, 2005). The outcome is the identification of groups of genes (modules) with similar expression behavior suggesting that they are under common transcriptional control. Modules (represented by module eigengenes) can be correlated with phenotypic variation to identify modules that may influence the trait (Li *et al.*, 2015).

A second systems genetics approach that can be implemented with transcriptome data is genetical genomics (Jansen and Nap, 2001), based on expression QTL (eQTL) analysis (Schadt *et al.*, 2003). This approach is essentially a massively parallel QTL mapping analysis in which the traits being mapped are the transcript abundance values for every gene that can be assayed using methods such as microarray or RNAseq analysis (Hansen *et al.*, 2008). eQTL analysis identifies genomic regions that are likely to contain causal polymorphisms with regulatory effects on the genes being assayed. Large-scale, global eQTL mapping studies on a variety of plants have been published over the past decade, including maize (Schadt *et al.*, 2003; Shi *et al.*, 2007; Swanson-Wagner *et al.*, 2009; Holloway *et al.*, 2011; Li *et al.*, 2013), Eucalyptus (Kirst *et al.*, 2005), Arabidopsis (Wentzell *et al.*, 2007; West *et al.*, 2007; Keurentjes *et al.*, 2007), wheat (Jordan *et al.*, 2007), barley (Potokina *et al.*, 2008; Druka *et al.*, 2008; Chen *et al.*, 2010; Moscou *et al.*, 2011), rice (Wang *et al.*, 2010; Wang *et al.*, 2014), cotton (Claverie *et al.*, 2012), potato (Kloosterman *et al.*, 2012) and *Populus* (Drost *et al.*, 2015). Several of these studies combined eQTL with QTL analysis to explore the genetic basis underlying phenotypic QTL, for example, Wentzell *et al.*, (2007) identified a candidate gene for glucosinolate accumulation in Arabidopsis.

A third systems genetics approach is to combine gene co-expression analysis with eQTL analysis to explore the genetic architecture of gene expression correlation in a population (Feltus, 2014). This was demonstrated using microarray data from embryos

of wheat plants grown in two different environments, which revealed conserved and environment-specific biological processes (Munkvold *et al.*, 2013). A fourth integrated systems genetics approach involves combining gene co-expression analysis, eQTL analysis and phenotypic analysis to prioritize candidate genes affecting trait variation in populations (Keurentjes *et al.*, 2007; Jiménez-Gómez *et al.*, 2010)

The aim of this study was to gain an understanding of the molecular basis of quantitative resistance and susceptibility to GLS in a RIL population of maize derived from a cross of sub-tropical inbred lines CML444 and SC Malawi (Berger *et al.*, 2014). We implemented an integrated systems genetics approach (Figure S1, Appendix S1) that combined transcript abundance variation, gene co-expression, eQTL and phenotypic QTL data to address the hypothesis that allelic variation gives rise to coordinated changes in gene expression, which in turn affect GLS disease severity.

RESULTS

Gray leaf spot disease severity and global gene expression profiling

We aimed to determine whether gene expression was correlated with GLS disease severity in a maize population segregating for quantitative resistance to GLS in the field. The CML444 X SC Malawi maize RIL population was planted at Baynesfield, South Africa, where *Cercospora zeina* is the causal agent of GLS disease (Meisel *et al.*, 2009). Typical GLS disease symptoms were observed, with some RILs exhibiting more extreme (transgressive) phenotypes than the resistant (CML444) and susceptible (SC Malawi) parental lines (Figure 1). GLS disease scores increased over time as the disease progressed in the field (Figure 1B). The distribution of disease scores per rating was characteristic of quantitative resistance (Figure 1B).

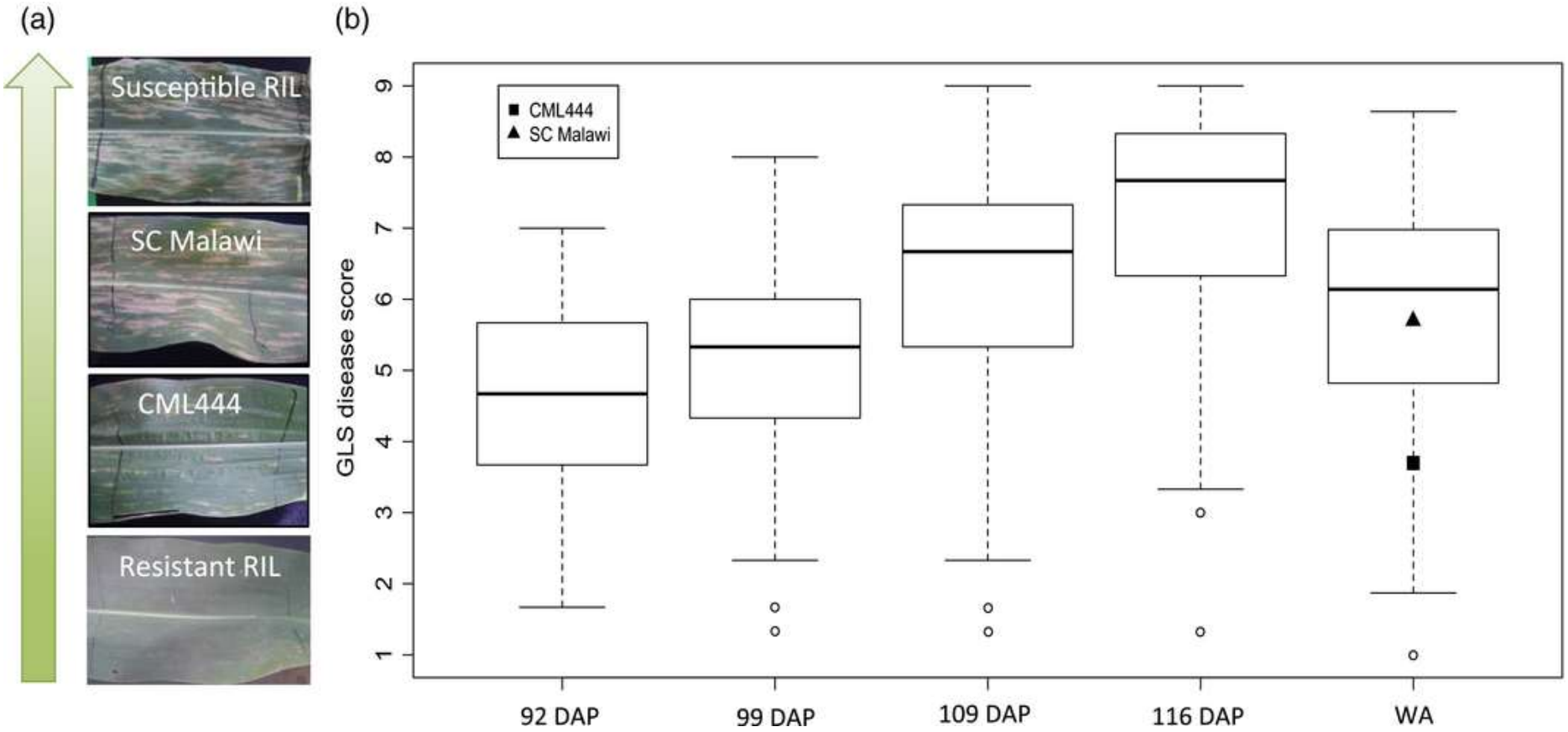

Figure 1. Grey leaf spot (GLS) disease responses in the CML444 × SC Malawi maize recombinant inbred line (RIL) population**.**

(a) Representative maize leaf samples collected for RNA extraction of a resistant RIL (GLS score = 1.5), parental line CML444 (GLS score = 3.7), SC Malawi (GLS score = 6.5) and susceptible RIL (GLS score = 8.5). GLS scores correspond to a scale from 1 to 9 shown on the *y*-axis of (b), and are based on whole-row plant scores as described in Berger *et al*. ([1]). **(b)** Boxplots of GLS disease severity data (*y*-axis), collected at 92, 99, 109 and 116 days after planting (DAP) at Baynesfield Estate in KwaZulu-Natal, South Africa. The right-hand boxplot represents a weighted average (WA) of the GLS severity scores across the four ratings, with weights based on the absolute value of the difference between the DAP at the rating and 103 DAP (when samples were collected for RNA analysis), with proportionally more weight given to ratings closest to 103 DAP. The GLS scores (WA) for CML444 and SC Malawi are indicated with a closed square and closed triangle, respectively.

Global gene expression profiling using Agilent 44K maize microarrays were carried out on earleaf samples collected from 100 RILs at 103 DAP. Expression profiles were obtained for 19,281 microarray reporters after filtering, normalization and back-conversion. As a first insight into the relationship between gene expression and disease, the least square means of a weighted average of the GLS severity scores across the four ratings (WA in Figure 1B) was compared with expression profiles of individual reporters across the RILs. We observed that 19% of reporters showed significant absolute correlations (False discovery rate (FDR)<0.05) with GLS disease scores (Table S1), indicating that co-expression may be relevant to GLS disease responses.

Identification of co-expression modules in the maize RIL population

To determine whether coordinated transcriptional responses to *C. zeina* infection were correlated with susceptibility and/or resistance in the maize RIL population, we conducted weighted gene co-expression network analysis (WGCNA) of the 100 RIL microarray data (Figure S1A). Approximately 50% of the reporters (8,665/19,281) were assigned to 42 co-expression modules, named after arbitrarily assigned colours (Table S2). Table 1 lists three of the modules (*GY-s*, *PT-s* and *TQ-r*) that were relevant to the GLS disease response, since the module eigengenes significantly correlated (FDR<0.05) to the GLS severity profile.

Gene Ontology enrichment and functional annotations of co-expression module GY-s

The *GY-s* module (185 reporters representing 179 genes; Table S3) had an exceptionally high positive correlation (0.71; FDR=4E-15, Table 1) with the GLS severity profile across the 100 RILs, i.e. higher expression was associated with increased susceptibility (Figure S2). This led us to hypothesize that the genes in this co-

Table 1. Co-expression modules from CML444 X SC Malawi maize RIL population with correlation to GLS disease scores, and comparison with genes expressed in susceptible maize inbred B73 during GLS disease

Co-expression module [a]	Module eigengene correlation with GLS severity	p-value (based on resampling)[b]	p-value (FDR)[c]	Number of reporters in module	Directional expression correlation with GLS severity[d]	% genes in module that are also induced in B73 - GLS [e]	Enrichment of module genes in induced gene list from B73 – GLS (FDR) [f]	% genes in module that are also repressed in B73 - GLS[g]	Enrichment of module genes in repressed gene list from B73 – GLS (FDR)[h]
GY-s	0.71	1E-16	4E-15	185	S	**55%**	**4E-57**	2%	ns
PT-s	0.31	0.003	0.03	41	S	15%	ns	10%	ns
TQ-r	-0.31	0.003	0.03	1564	R	9%	ns	**10%**	**0.049**

a Co-expression modules from WGCNA of microarray data of 100 RILs in the CML444 X SC Malawi RIL population exposed to GLS disease in the field.

b Null hypothesis: no correlation between the module eigengene expression values and GLS disease scores across RILs. P-values were calculated based on a resampling method (Methods S1).

c Calculated p-values were adjusted for multiple testing by controlling the false discovery rate (Benjamini and Hochberg, 1995).

d "S" indicates that co-expression module eigengene had a positive correlation with GLS severity scores, i.e higher expression values correlated with higher disease severity scores (susceptibility), and vice versa for "R" (resistance)

e Percentage of genes in each co-expression module of the CML444 X SC Malawi RIL population that were significantly induced (RNAseq; FDR<0.05) in susceptible maize inbred B73 plants infected with *C. zeina* compared to control material.

f Result of Fisher's Exact Test (adjusted p-value; FDR) for enrichment of module genes in list of genes induced in B73 infected with *C. zeina*. ns = not significant (i.e. p (FDR) >= 0.05)

g Percentage of genes in each co-expression module of the CML444 X SC Malawi RIL population that were significantly repressed (RNAseq; FDR<0.05) in susceptible maize inbred B73 plants infected with *C. zeina* compared to control material.

h Result of Fisher's Exact Test (adjusted p-value; FDR<0.05) for enrichment of module genes in list of genes repressed in B73 infected with *C. zeina*.

Table 2. Enriched GO-terms for the *GY-s* module *

GO description	FDR	Cluster freq	Total freq	GO-ID
diterpenoid/gibberellin metabolic process	1E-04	4/132 3.0%	5/8758 0.0%	16101/9685
catalytic activity	5E-04	73/132 55.3%	3091/8758 35.2%	3824
diterpenoid/gibberellin biosynthetic process	2E-03	3/132 2.2%	4/8758 0.0%	16102/9686
secondary metabolic process	3E-03	9/132 6.8%	101/8758 1.1%	19748
carboxylic acid / organic acid transmembrane transporter activity	5E-03	5/132 3.7%	27/8758 0.3%	46943/5342
cellular catabolic process	6E-03	13/132 9.8%	240/8758 2.7%	44248
small molecule metabolic process	6E-03	23/132 17.4%	641/8758 7.3%	44281
response to stimulus	6E-03	37/132 28.0%	1316/8758 15.0%	50896
carboxylic acid /organic acid transport	6E-03	5/132 3.7%	31/8758 0.3%	46942/15849
cellular nitrogen compound catabolic process	8E-03	4/132 3.0%	18/8758 0.2%	44270
lipid metabolic process	1E-02	13/132 9.8%	264/8758 3.0%	6629
heterocycle catabolic process	1E-02	4/132 3.0%	20/8758 0.2%	46700
monocarboxylic acid metabolic process	1E-02	9/132 6.8%	139/8758 1.5%	32787
gibberellic acid mediated signaling pathway	2E-02	3/132 2.2%	11/8758 0.1%	9740/71370/ 10476
malate/C4-dicarboxylate transmembrane transporter activity	3E-02	2/132 1.5%	3/8758 0.0%	15140/15556/ 5310
catabolic process	3E-02	13/132 9.8%	312/8758 3.5%	9056
small molecule biosynthetic process	3E-02	13/132 9.8%	313/8758 3.5%	44283
lipase activity	4E-02	4/132 3.0%	31/8758 0.3%	16298
carboxylic acid/oxoacid/organic acid metabolic process	4E-02	13/132 9.8%	324/8758 3.6%	19752/43436/ 6082
malate/C4-dicarboxylate transport/amino acid import	4E-02	2/132 1.5%	4/8758 0.0%	15743/15740/ 43090
cellular ketone metabolic process	4E-02	13/132 9.8%	331/8758 3.7%	42180
terpenoid metabolic process	4E-02	4/132 3.0%	35/8758 0.3%	6721
response to chitin	4E-02	4/132 3.0%	35/8758 0.3%	10200
aspartic-type endo/peptidase activity	5E-02	3/132 2.2%	17/8758 0.1%	4190/70001
response to organic substance	5E-02	14/132 10.6%	388/8758 4.4%	10033

* BiNGO was used to identify enriched GO-terms using ontology files downloaded from www.geneontology.org

expression module may represent targets of pathogen manipulation and/or represent orthologues of defence genes involved in other host-pathogen interactions. To this end we scrutinized protein, functional and GO term annotations of the reporters in the *GY-s* module.

The *GY-s* module was enriched for biological process in GO-terms diterpenoid/gibberellin biosynthesis, small molecule, lipid and secondary metabolism, and stress responses, such as peptidase activity and response to chitin (Table 2). The module included six reporters linked to calcium signaling including three calmodulin-like proteins with EF-hand calcium-binding motifs, three protein kinases, and several NAC and WRKY transcription factors (Table S3). Three components of jasmonate signalling (JAZ1, two different JAZ2 genes) were represented, as well as ethylene signaling (ACC synthase and an EIN3 binding protein).

Small molecule metabolism was reflected by reporters of two key enzymes of the glyoxylate pathway (isocitrate lyase and malate synthase) and fourteen reporters of lipid metabolism, including two acetyl-CoA synthases (Table S3). Five reporters had the GO term "carboxylic acid transport", and there were two SWEET sugar transporters. The module also included high proportions of reporters in the carbohydrate or energy metabolism, and protein or nucleotide catabolism categories.

Stress and defence responses were also represented in the *GY-s* module. These include a heat shock transcription factor, six ABC transporters, and two glutathione-S-transferases. Four reporters have the GO term "response to chitin" [WRKY and MYB transcription factors, a zinc finger (AN1-like) protein, and a fungal elicitor response protein]. Pathogenesis related (PR) and other defence proteins were also enriched in the *GY-s* module – four chitinases, two beta-glucanases, two osmotins, a viral response protein, and two protease inhibitors. Three key enzymes of diterpenoid/gibberellin

biosynthesis, ent-copalyl diphosphate synthase (An2), ent-kaurene synthase (KS2) and a P450 ent-kaurene oxidase were highly correlated with susceptibility in the *GY-s* module. Thirteen additional reporters involved in secondary metabolism were co-expressed. Taken together these results indicate that the *GY-s* module is made up of both genes involved in plant defence as well as genes with a role in nutrient acquisition which may represent the result of pathogen manipulation.

GO enrichment and functional annotations of GLS co-expression modules PT-s

Although the *PT-s* module (41 reporters; Table S4) had a lower and less significant correlation with GLS susceptibility (0.31, FDR=0.03), we wished to determine if there were any shared processes with the *GY-s* module. The *PT-s* module was enriched for the GO-term "ATP citrate synthase activity" (Table S4). Thus, the functional category primary metabolism appears to be the only common process between the *GY-s* and *PT-s* co-expression modules.

GO enrichment and functional annotations of GLS co-expression module TQ-r

In contrast to the previous two modules, the *TQ-r* module was made up of reporters with the opposite behavior across the RIL population, i.e higher expression in the more resistant RILs (Figure S2). This was reflected as a negative correlation with GLS disease scores (-0.31; FDR = 0.03)(Table 1). We therefore asked the question as to whether the *TQ-r* module (1564 reporters representing 1498 genes; Table S5) was characterized by genes and processes related to fungal disease resistance. The most relevant over-represented GO term was "cullin-RING ubiquitin ligase complex" (24 reporters)(Table S6), which is consistent with reports of the role of ubiquitination and proteolysis in plant defence responses (Furniss and Spoel, 2015). Interestingly, the reporter with the

highest *TQ-r* module membership score was annotated as coronatine-insensitive protein 1 (COI-1) (A_92_P001413, MM score = 0.94; Table S5). In addition, a COI-1 paralogue was co-expressed in the *TQ-r* module (A_92_P001595, Table S5). COI-1 is an F-box protein with roles in jasmonate signalling and regulation of NBS-LRR resistance protein accumulation and function, that has been well characterized in Arabidopsis (He *et al.*, 2012). Co-expression analysis also revealed the potential role of callose in resistance to GLS, since a callose synthase reporter (A_92_P010785) in the *TQ-r* module had the highest (negative) correlation to GLS disease severity (Table S5).

Genetic architecture of global gene expression and GLS disease phenotypic variation

Co-expression analysis had revealed groups of genes with intriguing annotations relevant to plant-pathogen interactions in the maize RIL population. However, the next aim was to explore the underlying genetics to determine whether there were genomic regions where allelic variation influenced gene expression and/or GLS disease phenotypes (Figure S1). To do this, we conducted expression QTL (eQTL) and GLS phenotypic QTL analyses of the CML444 X SC Malawi RIL population data from the Baynesfield trial.

Global eQTL analysis identified 20,206 eQTLs from the 19,281 input e-traits (Table S7). *Trans*-eQTLs accounted for 62% of the eQTLs, 22% were *cis*-eQTLs, and 16% had unknown classification (Figure S3, Table S8). *Cis*-eQTLs explained a significantly greater proportion of expression variation (average=30%) than *trans*-eQTLs (average=15%)(p=2E-16; Wilcoxon rank sum test)(Table S8).

Previous eQTL studies have identified genomic regions enriched for *trans*-eQTL (termed “hotspots”) that may represent genes subject to co-ordinated regulation (Li *et*

al., 2013). We aimed to find *trans*-eQTL hotspots that may be relevant to the co-expression modules and GLS disease responses. Twenty-one *trans*-eQTL hotspots were identified from the expression data of the maize RIL population exposed to GLS disease at 103 DAP (Figure S4; Table S9; see Methods S1 for permutation tests and calculations to discount effects of local gene density). A significant directional parental allelic bias was evident for 17 of the 21 *trans*-eQTL hotspots (Table S9), such that one of the parental alleles was associated with higher expression for most of the transcripts in a hotspot. The directional bias for the hotspots was in contrast to the global average for the 20,206 eQTLs, where 49% of had a positive CML444 allelic association and 51% had a positive SC Malawi allelic association.

QTL mapping of the maize RIL population was conducted using GLS disease scores from the same field trial from which RNA was sampled for global expression profiling (Figure S1C). Eight QTLs for GLS severity were identified (Table 3). Alleles from CML444 (the more resistant line) were associated with resistance (low disease severity) at six QTLs (QTL3a, QTL3b, QTL4, QTL5, QTL9a, QTL10), whereas alleles from SC Malawi were associated with increased resistance at only two loci, QTL6 and QTL9b (Table 3).

Having identified eight GLS QTLs and 21 *trans*-eQTL hotspots, the next step was to determine whether any of the *trans*-eQTL hotspots co-localized with GLS QTLs on the genetic map (Figure S1D). This would implicate the genes with *trans*-eQTL in the hotspot in processes leading to susceptibility or resistance to GLS conferred by the QTL. Four *trans*-eQTL hotspots were found to coincide with GLS severity QTLs (Table 4; Table S9). Permutation tests determined that obtaining four overlaps between 8 QTLs and 21 hotspots was significant (p=0.0045; See Methods S1 for calculations). The four hotspots coincided with GLS severity QTL4, QTL9a, QTL9b, and QTL10, and were

Table 3. GLS severity QTLs identified for the CML444 x SC Malawi maize RIL population

QTL name	Chromosome	Bin[a]	Peak marker[b]	1-LOD interval[c]	2-LOD interval[d]	LOD score[e]	R^2 [f]	Additive effect[g]	Allele source associated with Resistance [h]	Ratings[i]
QTL3a	3	3.03	bnlg1325	30.62 - 46.28	25.16 - 49.93	2.96	8.53	-0.48	CML	1,3*
QTL3b	3	3.08	umc16a	169.91 - 186.09	165.21 - 187.74	4.24	12.52	-0.47	CML	1,2*
QTL4	4	4.08	umc133a	115.16 - 127.68	93.53 - 131.68	3.06	7.01	-0.33	CML	1*
QTL5	5	5.02	umc90	26.96 - 43.53	25.61 - 47.81	2.94	7.87	-0.36	CML	1,2*
QTL6	6	6.06	umc1424_P	145.54 - 160.42	142.31 - 163.43	4.32	22.50	0.81	SC	3,4*
QTL9a	9	9.04	umc81	75.71 - 96.62	67.9 - 97.27	2.97	8.82	-0.37	CML	1*
QTL9b	9	9.06	umc1733	120.24 - 122.5	119.96 - 123.38	6.89	16.26	0.56	SC	1*,2,3
QTL10	10	10.07	bnl7.49a	113.93 - 127.28	110.84 - 129.74	4.56	17.22	-0.69	CML	3,4*

a Maize core bin that includes the start interval

b Peak marker refers to marker on QMap 2.0 that is closest to the QTL peak

c Range in cM that defines 1-LOD interval of QTL

d Range in cM that defines 2-LOD interval of QTL

e Log of odds (LOD) value at position of QTL peak

f Phenotypic variance explained by the QTL (expressed as percentage)

g Additive effect of QTL. Positive values indicate that the allele for resistance was derived from SC Malawi

h Parental allele associated with increased GLS resistance; CML = CML444; SC = SC Malawi

i GLS rating at Baynesfield Estate, KwaZulu-Natal (2008/2009 season) for which QTL was observed; 1, 2, 3, 4 refer to 92, 99, 109 and 116 days after planting; Stars indicate the rating with the highest LOD score (to which the statistics correspond)

* Rating that gave QTL with the highest LOD score that was selected for further analysis

Table 4. *Trans*-eQTL hotspots that overlap with GLS severity QTL in the maize RIL population, and co-expression modules enriched for reporters in these *trans*-eQTL hotspots

GLS QTL	GLS QTL 1-LOD interval (cM)[b]	*Trans*-eQTL hotspot name[c]	Chr	Hotspot interval (cM)[d]	eQTL[e]	eQTLs in subset[f]	allele[g]	*Trans*-eQTL hotspot subset name[h]	Bias[i]	*GY-s* (S)[j]	*TQ-r* (R)[k]	*Mg* (S)[j]	*Br* (R)[k]	*Mb* (S)[j]	*Pr* (S)[j]	*Yw* (R)[k]
QTL4	115.16 - 127.68	eQTL_HS4b	4	124.25 - 129.9	141	49 (35%)	CML	eQTL_HS4b(R)	yes							***
						92 (65%)	SC	eQTL_HS4b(S)							***	
QTL9a	75.71 - 96.62	eQTL_HS9a	9	92.13 - 101.81	407	135 (33%)	CML	eQTL_HS9a(R)	yes		***		***	***		
						272 (67%)	SC	eQTL_HS9a(S)		***						
QTL9b	120.24 - 122.5	eQTL_HS9b	9	117.81 - 124.51	142	88 (62%)	CML	eQTL_HS9b(S)	yes						***	
						54 (38%)	SC	eQTL_HS9b(R)			***					
QTL10	113.93 - 127.28	eQTL_HS10c	10	114.61 - 121.64	214	67 (31%)	CML	eQTL_HS10c(R)	yes				***			
						147 (69%)	SC	eQTL_HS10c(S)		***		***				

a QTL for GLS severity that overlaps with *trans*-eQTL hotspot (from Table 3)
b QTL interval on QMap 2.0 (Table 3)
c Name of *trans*-eQTL hotspot that overlaps with GLS severity QTL
d *Trans*-eQTL hotspot interval on QMap 2.0
e Total number of *trans*-eQTLs in each hotspot
f Number of *trans*-eQTLs in each hotspot divided into two subsets based on the allele that corresponds to higher expression
g Parental allele associated with higher expression
h *Trans*-eQTL hotspot subset name, based on the parental allele of the QTL associated with higher expression. For example for QTL4a, resistance is conferred by the CML444 allele (Table 3), and therefore 49 reporters have trans eQTL in the eQTL_HS4a hotspot where the CML444 allele (R) is associated with higher expression, whereas 92 reporters have the SC Malawi allele (S) associated with higher expression.
i Result of Pearson's chi-squared test to determine whether there was significant bias in the parental allele conferring higher expression for reporters within a trans eQTL hotspot ($P<0.05$).
j Module names: *GY-s*, *Mg* (Magenta), *Mb* (Midnightblue), *Pr* (Purple): "S" indicates that co-expression module had a positive correlation with GLS severity scores, i.e higher expression values correlated with higher disease severity scores (susceptibility)
k Module names: *TQ-r*, *Br* (Brown), *Yw* (Yellow): "R" indicates that co-expression module had a negative correlation with GLS severity scores, i.e higher expression values correlated with lower disease severity scores (resistance)
jk Enrichment of co-expression module reporter lists for reporters in trans eQTL hotspots was determined by Fishers Exact Test (*** FDR < 0.001)

termed eQTL_HS4b, eQTL_HS9a, eQTL_HS9b, and eQTL_HS10c, respectively. These four *trans*-eQTL hotspot regions spanned 4.4 cM to 9.7 cM and affected the expression of 141 to 407 reporters positioned throughout the genome (Table 4).

Since *cis*-eQTL have been shown in some cases to be causal of a phenotypic QTL (Drost *et al.*, 2015), we identified 129 *cis*-eQTL that overlapped the GLS phenotypic QTLs and correlated with GLS disease scores (FDR<0.05)(Table S10).

Genetic architecture of co-expression modules GY-s, PT-s* and *TQ-r

The final aim of the systems genetics strategy was to identify genomic regions that may contain polymorphisms leading to the downstream transcriptional responses reflected by the three GLS-significant co-expression modules (Figure S1E). First, we tested for significant enrichment of reporters from each co-expression module in the lists of reporters in each *trans*-eQTL hotspot, split by allele. Out of the three co-expression modules with a significant correlation to GLS severity (Table 1), two modules (*GY-s*, *TQ-r*) were significantly enriched for reporters with eQTLs in *trans*-eQTL hotspots that overlapped with GLS severity QTLs (Table 4; Table S9).

The *GY-s* module (highly correlated with GLS susceptibility) was enriched for reporters (see Table S11) with eQTLs in two hotspots (eQTL_HS9a(S) and eQTL_HS10c(S), Fisher's exact test, FDR<0.05; Table 4) that co-localized with GLS severity QTLs, namely QTL9a and QTL10, respectively. At both hotspot's, the parental allele associated with higher expression (SC Malawi) was also associated with higher susceptibility (Table 4).

Second, to identify genomic regions where genetic variation influences entire co-expression modules, we performed "*a posteriori* network eQTL analysis" (Hansen *et al.*, 2008), further referred to as "module eigengene eQTL analysis". The module eigengene

Table 5. Module eigengene eQTLs mapped for co-expression modules

Module name	Module eigengene correlation [a]	Chr	Peak position (cM)[b]	1-LOD interval[c]	2-LOD interval[d]	LOD score[e]	R^2 [f]	Additive effect [g]	Allele [h]	Overlapping GLS QTL
GY-s	S	9	96.14	80.75 - 103.8	71.95 - 109.46	2.74	9.18	-0.03	SC	QTL9a
GY-s	S	10	117.64	111.89 - 126.13	107.48 - 129.65	2.83	9.03	-0.03	SC	QTL10
PT-s	S	1	38.72	33.39 - 42.83	28.13 - 45.88	3.38	11.91	0.04	CML	
PT-s	S	2	132.63	131.97 - 136.07	131.3 - 138.88	3.71	13.20	0.05	CML	
TQ-r	R	7	55.45	51.05 - 59.09	47.6 - 62.45	3.04	11.51	-0.04	SC	

a "S" indicates that co-expression module had a positive correlation with GLS severity scores and "R" indicates that co-expression module had a negative correlation with GLS severity scores

b Position of the module eigengene eQTL peak in cM

c Range in cM that defines 1-LOD interval of eQTL

d Range in cM that defines 2-LOD interval of eQTL

e Log of odds (LOD) value at position of eQTL peak. The *GY-s* module eigengene on chromosome 9 was retained since it was close to the threshold of LOD=2.8.

f Phenotypic variance explained by the eQTL (expressed as percentage)

g Additive effect of eQTL. Positive values indicate that the allele associated with higher expression was derived from CML444

h Parental allele associated with higher expression; CML = CML444; SC = SC Malawi

profiles of the gene co-expression modules were used as the traits in a QTL analysis (Table 5). For the *GY-s* module, two module eigengene eQTLs co-localized with the GLS phenotypic QTL9a (and eQTL_HS9a(S)) and QTL10 (and eQTL_HS10c(S)), respectively (Table 5). In both cases, the SC Malawi parental allele had higher expression and was associated with increased susceptibility. This corroborated the first meta-analysis that the reporters in the *GY-s* module had a significant number of eQTLs in *trans*-eQTL hotspots eQTL_HS9a(S) and eQTL_HS10c(S) (Table 4).

Figure 2 is graphical representation of the *GY-s* module illustrating the overlap between phenotypic QTL9a or QTL10, *trans*-eQTL hotspots, and *trans*-eQTLs of reporters in the module. Enrichment of *GY-s* reporters with *trans*-eQTL in hotspot eQTL_HS9a can be seen by the many blue links (track f) emanating from the dark blue line (hotspot eQTL_HS9a, track e; corresponding to 92-102 cM on chromosome 9, track a) that coincides with QTL9a (dark orange block, track c). Likewise, enrichment of *GY-s* reporters with *trans*-eQTL in the hotspot eQTL_HS10c can be seen by the many blue links (track f) emanating from the dark blue line (hotspot eQTL_HS10c, track e; corresponding to 114-121 cM on chromosome 10, track a) that coincides with QTL10 (dark orange block, track c). Two *GY-s* module eigengene eQTLs are shown as greenyellow bars in track (d) of the circos plot, and their positions correspond to GLS severity QTL9a and QTL10, respectively (dark orange blocks in track (c) of Figure 2).

The *PT-s* module (correlated with susceptibility) was not enriched for reporters with *trans*-eQTLs in any of the *trans*-eQTL hotspots (Table S9). Two module eigengene eQTL were mapped for this module (on chromosome 1 and 2), however these did not coincide with GLS QTLs (Table 5), and therefore module *PT-s* was not investigated further in detail.

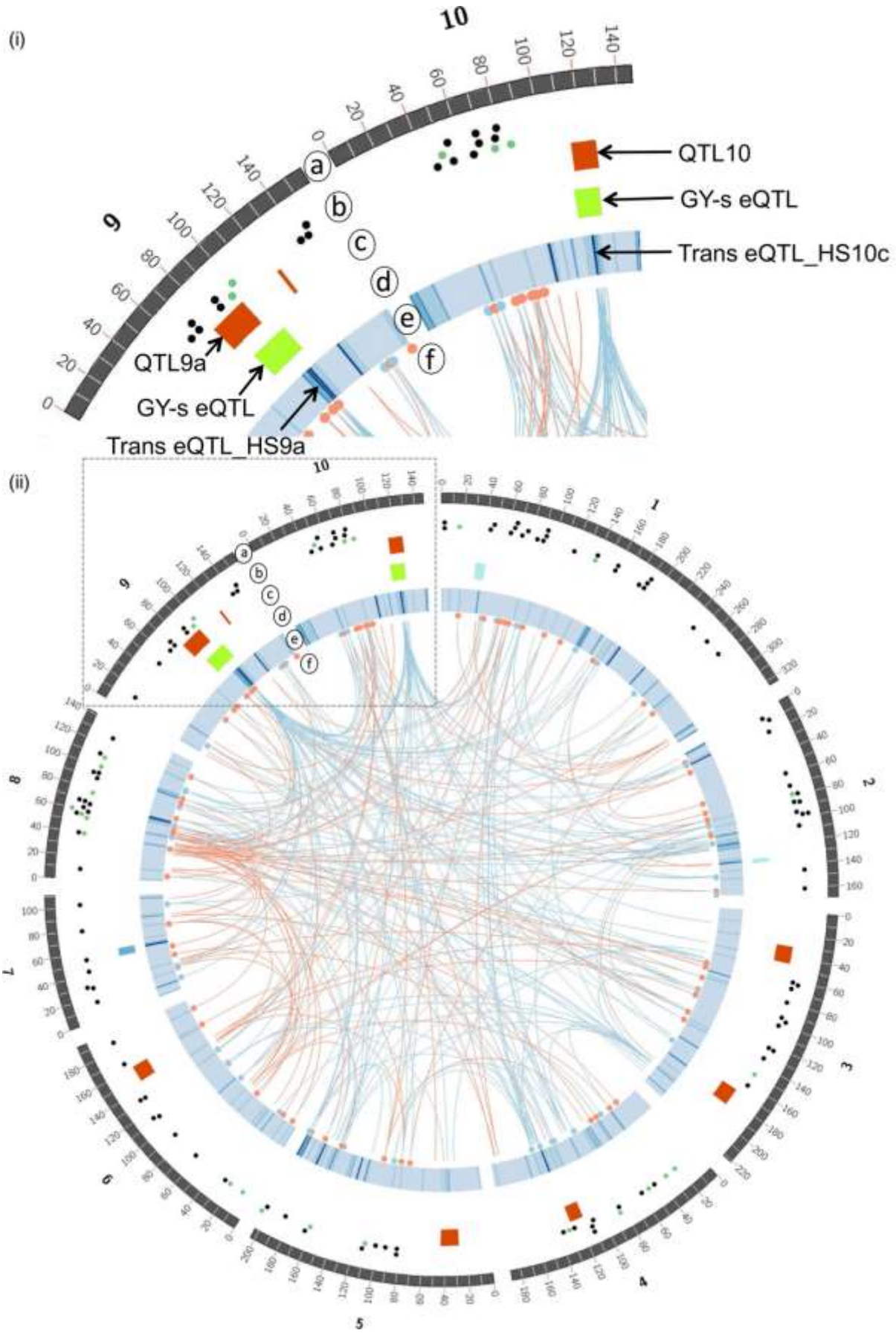
(i)
10
9
QTL10
GY-s eQTL
Trans eQTL_HS10c
QTL9a
GY-s eQTL
Trans eQTL_HS9a
(ii)

Figure 2. Genome view of transcriptional regulation underlying the grey leaf spot (GLS) susceptibility-associated *GY-s* co-expression module in relation to GLS quantitative trait loci (QTLs) and expression QTLs (QTL*s*).

(i) Detail of chromosomes 9 and 10 corresponding to the stippled box from (ii).

(ii) The whole genome. (a) Chromosomes: the white lines represent 10-cM bins with the cM position for every second bin shown on the outside. (b) Reporters: the dots (black and green) represent the 185 reporters in the *GY-s* co-expression module (Table S3). Green dots represent reporters with *cis*-eQTL*s*. (c) QTL: the dark orange blocks represent the eight GLS severity QTLs (Table 3). (d) Module eigengene eQTL: each colour bar represents module eigengene eQTL*s* for the co-expression modules that had significant absolute correlation with GLS severity scores, i.e. *GY-s* on chromosomes 9 and 10, *PT-s* on chromosomes 1 and 2, and *TQ-r* on chromosome 7 (Table 5). (e) Global *trans*-eQTL density heatmap. Dark blue indicates *trans*-eQTL hotspots (Tables S7 and S9). (f) Links: *trans*-eQTL*s* of reporters in the *GY-s* co-expression module. A dot at the end of a link represents the position of a reporter and nothing at the end of a link represents the position of an eQTL. Red or blue links connect reporters with *trans*-eQTL*s* for which the parental allele associated with higher expression was CML444 or SC Malawi, respectively.

The *TQ-r* module was enriched for reporters with eQTLs in eight hotspots (Table S9), of which two overlapped GLS severity QTLs (eQTL_HS9a(R) and eQTL_HS9b(R); Table 4; see Table S12 for list of reporters). In both cases the parental allele with increased expression was positively associated with resistance (Table 4). For the *TQ-r* module, only one module eigengene eQTL was identified (on chromosome 7), but this did not overlap with a GLS severity QTL (Table 5; Figure 2, track d). However, it overlapped one of the *trans*-eQTL hotspots for which the *TQ-r* module was strongly enriched (*trans*-eQTL hotspot eQTL_HS7, adj. p-value = 8.E-26; Table S8). These analyses provided some evidence for allelic variation at GLS QTLs underlying co-expression of genes in the *TQ-r* module.

Co-expression patterns in maize B73 challenged with C. zeina

Co-expression analysis of the sub-tropical maize RIL population had revealed the *GY-s* module as the only one with a highly significant correlation with GLS disease scores (r=0.71, FDR=4E-15; Table 1). We therefore focused our attention on the susceptible response and asked the question whether any of the co-expression modules were enriched in the GLS susceptible response of maize inbred line B73.

An independent field experiment at Hildesheim, KwaZulu-Natal was conducted with B73 challenged with *C. zeina*. B73 is susceptible to GLS caused by *C. zeina*, therefore typical GLS disease lesions formed on lower leaves of B73 maize plants prior to sampling of leaves at 77 DAP for RNA extraction and RNAseq analysis (Methods S1).

RNAseq gene expression data from B73 was obtained for 14,342 maize genes, and 2368 and 2079 genes were up- and down- regulated in response to *C. zeina*, respectively (FDR<0.05)(Table S13, Methods S1). Most notably, 55% of the reporters in

the *GY-s* module were also significantly up-regulated in B73 in response to *C. zeina* (Fisher's exact test, FDR=6E-59)(Table 1). Enriched GO terms in the 99 shared genes (corresponding to 103 microarray reporters) were lipid metabolite process and small molecule metabolite process (including key enzymes in the glyoxylate pathway – isocitrate lyase and malate synthase; Table S14.1), which were also enriched in the *GY-s* module (Table 2). This "common susceptible response" also included genes encoding enzymes in diterpenoid/gibberellin biosynthesis, PR proteins, stress/detoxification, protein catabolism, hormone responses, secondary metabolism, signaling and transcription (Table S14.2, Table S3). Interestingly, 59% of the (58) reporters in the *GY-s* module with eQTLs that coincided with QTL9a or QTL10 correspond to genes that were also found to be induced in the B73-*C. zeina* dataset (Table S11).

The *TQ-r* and four other co-expression modules not correlated with GLS disease in the RIL population (blue, royalblue, brown, tan; Table S15) showed enrichment of genes in the B73 susceptible response to *C. zeina* (FDR<0.05; Table 1), however they did not exhibit the high percentage of enrichment shown for the *GY-s* module (Table 1).

From these results, we conclude that the overlap of 99 genes in the *GY-s* module and the B73-*C. zeina* data (i) represents a common susceptible transcriptional response; and (ii) served to technically validate the microarray data in the RIL population using an independent technique (RNAseq).

Kauralexins are induced in response to *C. zeina* in maize line B73.

GO terms related to diterpenoid/gibberellin biosynthesis were enriched in the *GY-s* module, since reporters encoding three enzymes of the biosynthesis pathway (ent-copalyl diphosphate synthase (An2), kaurine synthase (KS2) and a p450 kaurine

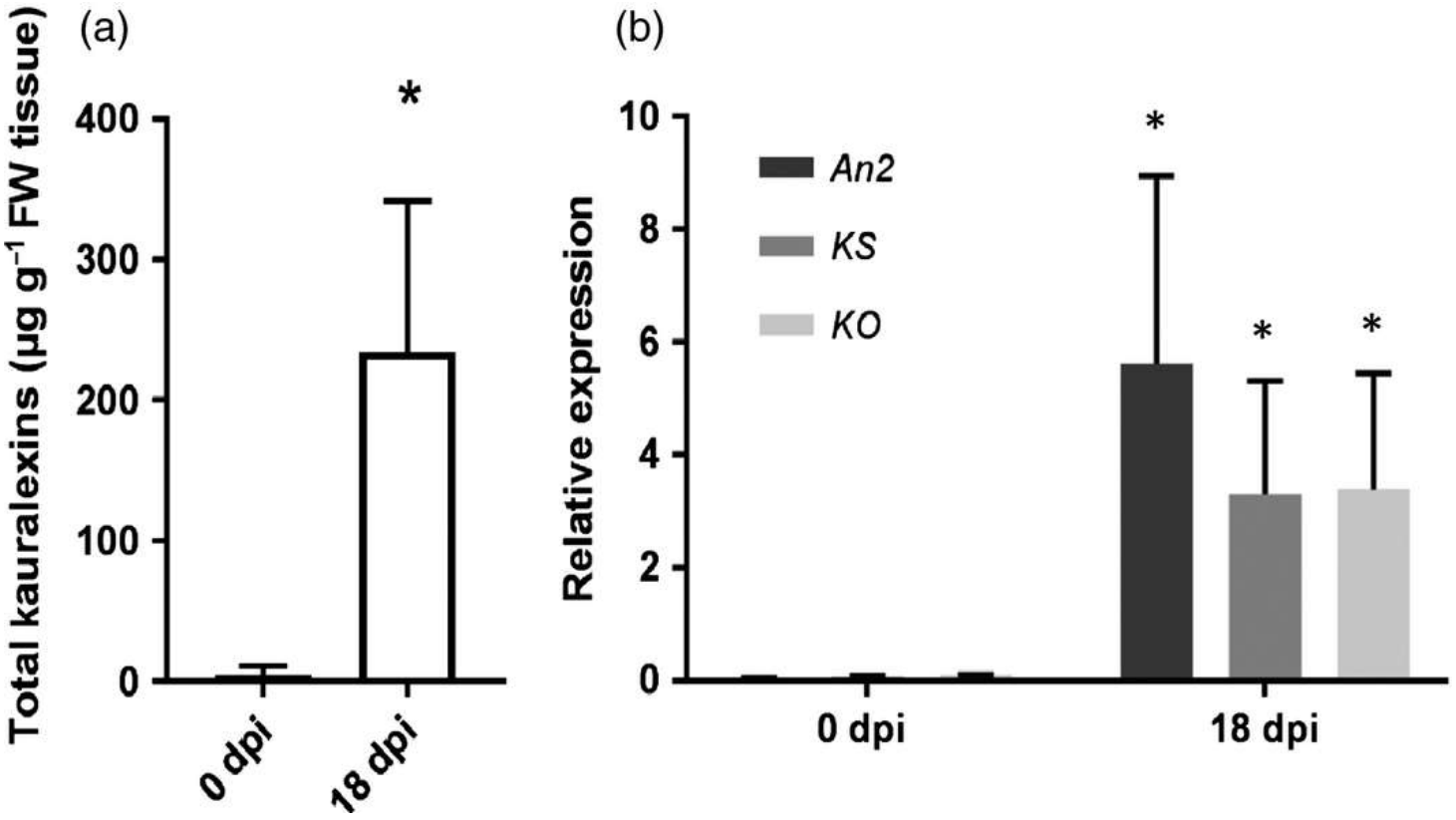

Figure 3. Kauralexins are induced by *Cercospora zeina* infection in susceptible maize line B73.

(a) Analysis of total kauralexin accumulation and **(b)** relative expression of putative kauralexin biosynthetic genes ent-copalyl diphosphate synthase (*An2*), ent-kaurene synthase (*KS2*) and a P450 ent-kaurene oxidase (*KO*), at 0 and 18 days post-inoculation (dpi) following inoculation of B73 plants with *C. zeina* GC-MS was used to determine the accumulation of total kauralexins.

(b) RT-qPCR was used to determine relative expression of *An2*, *KS* and *KO*, normalised to reference genes *Lugein* and *RNA polymerase II*. In all cases the means of three samples are shown and bars on the graph indicate standard deviation. Asterisks above the graph indicate significantly different accumulation or expression in 18-dpi samples compared with 0 dpi (unpaired *t*-test; $P < 0.05$).

oxidase) were co-expressed in the sub-tropical maize population (Table S3). Furthermore, these genes were induced 9-10 fold in *C. zeina* infected maize line B73 (Table S13). This raised the question as to whether co-expression of the pathway was the result of fungal manipulation of gibberellin biosynthesis to support disease or phytoalexin biosynthesis by the maize host.

Maize B73 plants were inoculated with *C. zeina* in a glasshouse experiment, and samples collected prior to inoculation (0 dpi) and after the development of GLS disease lesions (18 dpi). RT-qPCR confirmed induction of all three diterpenoid biosynthesis pathway genes and chemical analysis showed accumulation of total kauralexins in response to *C. zeina* (Figure 3). Kauralexins have anti-fungal activity (Schmelz *et al.*, 2011), and thus expression of the pathway is most likely a maize host defence response, albeit unsuccessful.

DISCUSSION

The systems genetics approach employed in this study enabled detailed exploration of the genetic architecture of gene expression in a sub-tropical maize RIL population (CML444 x SC Malawi) grown in the field under GLS disease pressure. We utilized a range of filters to discriminate between responses to *C. zeina* challenge and other foliar traits. First, gene co-expression network analysis revealed that three out of the 42 modules were correlated with GLS disease scores. Second, eQTL analysis identified four *trans*-eQTL hotspots that coincided with phenotypic QTLs for GLS severity. Third, meta-analysis of co-expression modules, eQTLs, and phenotypic QTL data allowed identification of genomic regions harboring polymorphisms contributing to co-expression modules *GY-s* and *TQ-r* as well as trait variation. Finally, we uncovered a

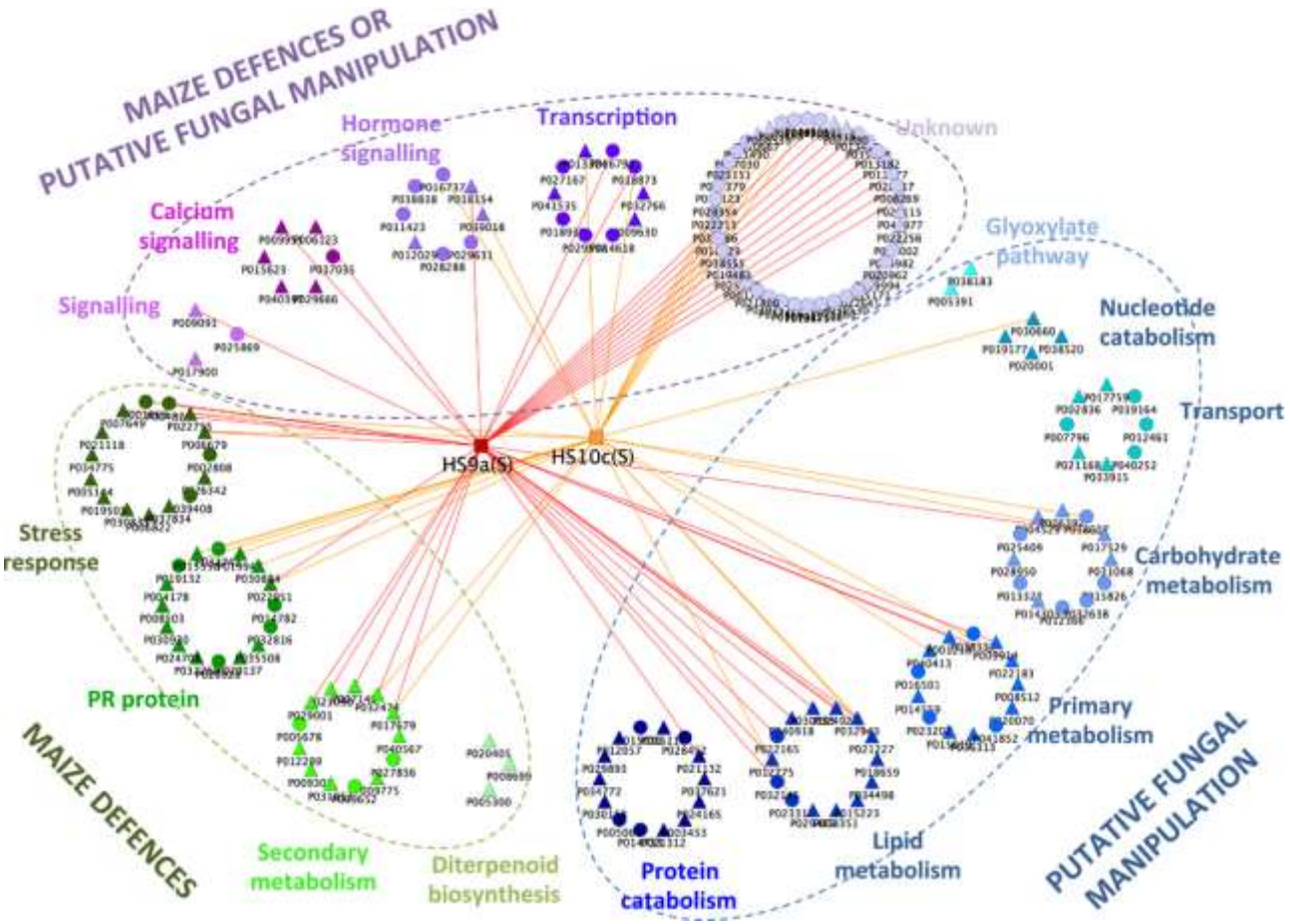

Figure 4. Model of the transcriptional network associated with susceptibility based on the *GY-s* co-expression module and genes induced in maize line B73 by *Cercospora zeina*.

Reporters are indicated by filled spheres or filled triangles (if the corresponding gene is also induced in B73 by *C. zeina*), and grouped according to functional categories noted in Table S3. Functional categories were further grouped based on whether their expression is proposed to be the result of putative fungal manipulation, defence responses or either. Reporters with *trans*-expression quantitative trait loci in hotspots 9a(S) or 10c(S) are shown with lines emanating from boxes labelled HS9a(S) or HS10c(S), respectively.

subset of 99 genes in the susceptibility-associated *GY-s* module that were up-regulated in the susceptible maize inbred line B73 in response to GLS measured with an independent technique, RNAseq.

Many QTLs for GLS disease severity have been mapped in maize (Berger *et al.*, 2014; Wisser *et al.*, 2006), however very little is known of the molecular responses downstream of the recognition event leading to resistance or susceptibility. This study uncovered a transcriptional network based on the *GY-s* co-expression module of 185 reporters (representing 179 genes) that were highly correlated with GLS disease scores across the sub-tropical maize population. Based on the functional processes and genes that were co-expressed, we propose a model of how the network may be the product of both manipulation by the fungus and maize defence responses (Figure 4).

First, a suite of processes related to nutrient acquisition are co-expressed, particularly the glyoxylate pathway, dicarboxylic acid transport and lipid metabolism, which points to the conversion of lipid derived acetate into energy rich carbohydrates (Figure 4). Genes involved in lipid metabolism and elements of the glyoxylate pathway (citrate synthase and isocitrate lyase) were shown to be induced during the necrotrophic phase of the wheat leaf pathogen *Zymoseptoria tritici* (Rudd *et al.*, 2015). In this study, two SWEET sugar transporters were co-expressed, and several components of primary metabolism, as well as protein and nucleotide catabolism (Figure 4). Sugar uptake has been shown to be important for virulence of other maize pathogens, namely *Ustilago maydis* and during the hemibiotrophic phase of *Colletotrichum graminicola* (Lingner *et al.*, 2011; Schirawski, 2015). A noteworthy maize gene in the network is an orthologue of At_WRKY75 involved in protection of Arabidopsis plants from phosphorus starvation (Devaiah *et al.*, 2007).

Second, genes encoding maize defence responses are co-expressed in the network, namely three enzymes associated with maize kauralexin biosynthesis, 13 genes involved in secondary metabolism, several genes encoding PR proteins (chitinases, beta-1-3-glucanases, osmotins) or with the GO term "response to chitin" (Figure 4), including a MYB transcription factor and an orthologue of AtWRKY53, a positive regulator of basal defence responses (Murray *et al.*, 2007). PR protein and secondary metabolite pathway gene expression was also observed in susceptible wheat leaves infected with *Zymoseptoria tritici* (Rudd *et al.*, 2015). Expression of PR proteins was reported in the interaction between *C. graminicola* and a highly susceptible maize line (Vargas *et al.*, 2012).

Stress responses, such as chaperones, ABC transporters and glutathione-S-transferases were also represented in the maize susceptibility network (Figure 4), illustrating the importance of detoxification in the response to *C. zeina* infection. The phytotoxin cercosporin is considered to be a hallmark of GLS disease (Shim and Dunkle, 2002). Wisser *et al.*, 2011 identified SNPs in a maize glutathione S-transferase gene explaining variation in response to GLS in a panel of diverse inbred lines by association mapping. In our study, although the reporter (A-92_P031881) for this gene did not cluster with any of the co-expression modules, its expression was significantly correlated with low GLS scores (Table S1), and it had a large effect cis-eQTL (LOD = 17.9) with higher expression associated with the CML444 allele (Table S7).

Third, a set of processes that could fall into either category of responses was included in the network, namely signaling and transcription (Figure 4). Calcium signaling is represented by two calmodulin-like proteins with triple EF-hand Ca+2 binding motifs, a calmodulin-binding protein, a calcium dependent protein kinase, and a calcium-

dependent lipid binding protein. Calcium signaling is central to plant immunity, and in Arabidopsis, calmodulin-like calcium binding proteins that are positive and negative regulators of defence gene transcription have been reported (Seybold *et al.*, 2014).

In Arabidopsis, jasmonate signaling leads to resistance to necrotrophic pathogens (Tsuda and Somssich, 2015), and the JASMONATE ZIM domain proteins (JAZ) are repressors of jasmonate response gene expression (Thines *et al.*, 2007; Zhang *et al.*, 2015). In our study, maize orthologues of Arabidopsis JAZ1 and two JAZ2-like proteins were co-expressed in the *GY-s* module (hormone signaling category, Figure 4), raising the intriguing possibility that they may be targets of effectors to promote susceptibility, as has been shown previously in Arabidopsis (Jiang *et al.*, 2013). Another maize gene in the *GY-s* module was an orthologue of the rice and Arabidopsis CHY RING zinc-finger domain proteins, the mutants of which show reduced stomatal aperture (Hsu *et al.*, 2014). GLS pathogens enter maize leaves through stomata (Kim *et al.*, 2011), thus this gene may represent another target of the pathogen to facilitate stomatal entry.

The model in Figure 4 illustrates that components of the susceptibility transcriptional network share *trans*-eQTLs in the hotspots that coincide with QTL9a and QTL10, which may indicate gene(s) that influence gene expression in the network (Table S11). Parental line SC Malawi alleles at *trans*-eQTL hotspot HS9a(S) result in higher expression of two calcium signaling genes (Figure 4). Interestingly, the only reporter in the network with a *cis*-eQTL that co-localizes with the susceptible allele of QTL9a and its *trans*-eQTL hotspot encodes a calmodulin-like protein (A_92_P037035) (Table S10). Figure 4 illustrates that neither *trans*-eQTL hotspots are predominantly associated with responses grouped as fungal manipulation or maize defences, although *trans*-eQTL HS9a(S) is associated with eight reporters of lipid metabolism. The three

diterpenoid (kauralexin) biosynthesis reporters do not have *trans*-eQTLs in these hotspots, however they each have overlapping *trans*-eQTLs on chromosome 8 which may represent a site of co-regulation (Table S7).

Finally, a significant proportion of genes in the network (55%) from the sub-tropical maize RIL population are induced by *C. zeina* in the susceptible temperate model maize inbred B73 (reporters corresponding to these genes indicated by triangles in Figure 4). These include both putative fungal manipulation targets, such as the glyoxylate pathway, lipid metabolism, dicarboxylic acid transport, and catabolic processes, as well as defense responses such as stress responses, PR proteins and secondary metabolism, including the three enzymes of diterpenoid biosynthesis. In this regard, metabolite analysis showed a significant increase in the products of this pathway (i.e. kauralexins) in *C. zeina* infected B73 plants (Figure 3). This result supports the hypothesis that induction of this pathway represents maize defence responses, since kauralexins have anti-fungal activity (Schmelz *et al.*, 2011), rather than fungal manipulation of gibberellin biosynthesis. A reason for the lack of effectiveness of kauralexins may be that *C. zeina* is able to detoxify, tolerate or even utilize kauralexins, similar to *Gibberella pulicaris* which detoxifies phytoalexins to enhance virulence on potato tubers (Desjardins *et al.*, 1992). Genes involved in calcium signaling, transcription and hormone signaling are also common to the B73 susceptible response. Taken together, this work has uncovered a transcriptional network associated with susceptibility to *C. zeina* in maize made up of candidate genes that may be targets of fungal manipulation, and defence response genes. Expression of the latter group can be explained by two hypotheses, (i) anti-fungal products that are ineffective at a late stage of pathogen establishment: or (ii) inappropriate “defences” that do not target *C. zeina*, in an opposite fashion to

Pseudomonas syringae that promotes jasmonate signaling to suppress anti-bacterial salicylic acid signaling (Zhang *et al.*, 2015).

This study also identified a transcriptional network associated with resistance to GLS in maize, based on the *TQ-r* co-expression module, which may represent downstream events after recognition of the fungus in plants with resistance alleles on chromosome 9 (QTL9a and QTL9b). This module had lower absolute correlation with GLS disease scores than the *GY-s* susceptibility module which may reflect the time of sampling (103 DAP). Resistant RILs may have mounted induced responses earlier in the season, which may have been dampened by 103 DAP, although GLS is a polycylic disease and therefore continuous inoculum pressure would be expected (Ward *et al.*, 1999). Nevertheless, suggestive resistance-related candidate genes were found to be expressed in the *TQ-r* network.

A maize orthologue of Arabidopsis COI-1, the jasmonate receptor, had the highest module membership in the *TQ-r* network, which implicates jasmonate responses in resistance to GLS. This fits the model that jasmomate signaling confers resistance to necrotrophs (and the necrotrophic phase of pathogens such as *C. zeina*) (Glazebrook, 2005). Further, our data fit the current biochemical model of jasmonate-induced gene expression described in Arabidopsis (Zhang *et al.*, 2015), namely positive regulators (COI-1 and a paralogue) are co-expressed in the resistant module *TQ-r*, and negative regulators (three JAZ genes) are co-expressed in the susceptible module *GY-s*. Furthermore, in Arabidopsis, COI-1 is the F-box subunit of an SCF-type ubiquitin E3 ligase. The *TQ-r* module was enriched for the GO term "cullin-RING ubiquitin ligase complex" that included a quarter of all genes with this term, encoding an E3 ubiquitin ligase SCF-complex subunit SKP1, a SKP1-interacting protein, a SKP1-like protein, a F-box protein, and 15 proteins with transducin/WD40 domains. These include

orthologues of MAX2 and SCD1 of Arabidopsis, both of which are involved in stomatal aperture and defence responses (Korasick *et al.*, 2010; Piisilä *et al.*, 2015)(Table S6). Taken together, co-expression of these genes in the *TQ-r* network implicate ubiquitination and protein turnover in resistance to *C. zeina*, possibly via COI-1 and jasmonate responses (Nagels Durand *et al.*, 2016).

Callose deposition in the form of local cell wall thickenings, called papillae, is a typical response of plants to fungal attack (Ellinger *et al.*, 2013). Hinch and Clarke (1982) documented callose formation in maize roots as a response to infection with *Phytophthora cinnamomi*. Transgenic Arabidopsis expressing a callose synthase *pmr4* had elevated early callose deposition that halted penetration of the powdery mildew pathogen (Ellinger *et al.*, 2013). More recently, PMR4 was shown to be an effector of a RabA4c GTPase required for callose biosynthesis (Ellinger *et al.*, 2014). In our study, expression of a maize callose synthase was the most highly correlated with low GLS disease scores in the *TQ-r* module, and the module also contained a Rab GTPase (A_92_P009206). Although these are not predicted to be the maize orthologues of *pmr4* and it's GTPase partner, these results implicate callose biosynthesis in resistance conferred through the TQ-r module. The orthologue of Arabidopsis *pmr4* (A_92_P000970) together with three other callose synthases (A_92_P001619; A_92_P016646; A_92_P005832) (Table S1) had expression correlation with resistance (i.e. low disease scores), but they were not included in *TQ-r*.

A limitation of the experimental design of our study was that we were unable to distinguish between constitutive or induced responses to *C. zeina* in the maize RIL population, since we captured both responses by sampling plants challenged with GLS disease in the field, and did not replicate the whole trial with fungicide treatments. However, allelic differences in both types of responses could be relevant to resistance or

susceptibility to GLS. For example, a wheat lipid transfer protein was expressed 50-fold higher in the absence of pathogen in a line carrying a QTL for *Fusarium* resistance compared to the isogenic line lacking the QTL (Schweiger *et al.*, 2013).

In addition to the three co-expression modules correlated with GLS disease responses, we considered whether our data could reveal biological insights for other foliar traits. Maize inbred line CML444 exhibits improved drought tolerance traits compared to SC Malawi, namely higher chlorophyll content, smaller leaves, higher photosynthetic capacity under water stress, and deeper roots (Messmer *et al.*, 2009; Trachsel *et al.*, 2009; Trachsel *et al.*, 2010). QTLs have been mapped for those traits in this population (Trachsel *et al.*, 2010), and some of the leaf co-expression modules unrelated to GLS could reflect such allelic differences. Interestingly, two of the *trans*-eQTL hotspots identified in our study coincided with QTLs for early vigour traits described by (Trachsel *et al.*, 2010), namely (i) *trans*-eQTL HS_5c (Table S9) which coincided with the QTL for leaf area and shoot dry weight in bin 5.08, and (ii) *trans*-eQTL_HS_7 (Table S9) which coincided with a QTL for leaf chlorophyll content in bin 7.03.

Expression QTL analysis at the level of resolution provided by the CML444 X SC Malawi RIL population of 100 individuals and a genetic map with 167 markers revealed 20,206 eQTLs for 12,725 reporters out of 19,281 reporters tested. The number of eQTLs detected per reporter (1.6) is consistent with previous eQTL studies in maize which ranged from 1.1 – 3 (Schadt *et al.*, 2003; Swanson-Wagner *et al.*, 2009; Li *et al.*, 2013; Shi *et al.*, 2007). Previous authors noted that the effect of *cis*-eQTLs were significantly greater than that of *trans*-eQTLs (Holloway *et al.*, 2011; Li *et al.*, 2013; Schadt *et al.*, 2003; Swanson-Wagner *et al.*, 2009), and we observed the same in our data with *cis*-eQTLs explaining on average a greater percentage of the variation in

expression than *trans*-eQTLs (30% vs 15%, respectively). Several explanations have been put forward for the differences in the effect of each type of eQTL, for example (i) that major effect *trans*-eQTLs may be deleterious (Swanson-Wagner *et al.*, 2009), or (ii) it is the result of ascertainment bias between *cis*-eQTL (only one possible locus) and trans-eQTL (multiple loci across the genome).

We observed a striking pattern of parental allelic bias for 17 of the 21 *trans*-eQTL hotspots (Table S9), an observation also made in several previous eQTL studies (Li *et al.*, 2013; Swanson-Wagner *et al.*, 2009). An explanation for this could be that one or more polymorphisms underlie the *trans*-eQTL hotspot with a particular parental allele or haplotype causing higher expression of the majority of the target genes with eQTLs in the hotspot.

In conclusion, the identification of a transcriptional network associated with susceptibility to the pathogen causing GLS disease in maize that has elements in common between sub-tropical and temperate maize highlights potential targets of *C. zeina* effectors. Our data also raises the hypothesis that the fungus exploits the host's glyoxylate pathway and lipid metabolism to release energy rich carbohydrates. Involvement of the jasmonate pathway in resistance to *C. zeina* is supported by expression of the jasmonate receptor COI-1 and enrichment of ubiquitination processes in resistant plants, and co-expression of JAZ repressors in susceptible plants. Further exploration of these hypotheses will require development of near-isogenic lines, fine-mapping of large-effect eQTL and mutant/over-expression studies of candidate genes.

EXPERIMENTAL PROCEDURES

Germplasm and field trials

A recombinant inbred line population (145 RIL, F7:S6) derived from a cross between subtropical white dent maize inbred lines CML444 and SC Malawi (Messmer *et al.*, 2009) was planted at Baynesfield Estate in KwaZulu-Natal Province, South Africa. GLS disease severity data was collected at 92, 99, 109 and 116 days after planting (DAP) (Figure 1). For the validation experiment, maize inbred line B73 was planted in three replicate rows of 10 plants each at Hildesheim Research Station, PANNAR SEED Pty Ltd, Greytown, KwaZulu-Natal, South Africa (Methods S1).

RNA extraction and microarray analysis

RNA was extracted from leaf samples of three biological replicates of 100 selected CML444 X SC Malawi maize RILs, sampled at Baynesfield at 103 DAP. Microarray expression data was obtained for 100 RILs after two-colour Cy-dye labeling of cDNA and hybridization to Agilent-016047 maize 4×44 K microarrays (Coetzer *et al.*, 2011) (Methods S1). Microarray data have been deposited in the NCBI Gene Expression Omnibus (Accession # GSE76242).

Gene co-expression network analysis

Weighted gene co-expression network analysis (WGCNA) was performed as described previously (Langfelder and Horvath, 2008) with an R-script modified for this study (Methods S1, Methods S2). The input data matrix consisted of expression values of 30,280 microarray reporters for each of the 100 RILs. After normalization and filtering 19,281 reporters remained in the data set, representing 14,201 maize gene

models (Coetzer *et al.*, 2011). The output of WGCNA was the identification of co-expression modules and calculation of an "eigengene" for each module (a summary expression profile representing the module) (Langfelder and Horvath, 2008).

A weighted average GLS disease score (WA) was calculated for the three replicates of each RIL. The weight for each of the four rating days were calculated as follows: ((total number of days that span the sampling period (24)) – (the absolute value of the difference between 103 DAP (RNA sampling day) and DAP for the rating))/(total number of days that span the sampling period (24)), expressed as a proportion of the sum across the four ratings. The weighted average for each RIL replicate was then calculated by multiplication of the score at each rating by its weight, summed across the four ratings. Thus, more weight was given to ratings that were closer to 103 DAP, the time point when samples were collected for RNA analysis (Figure 1). The least square means of the weighted average GLS disease scores were used for correlation analysis with expression profiles. Pearson correlation was used to determine the correlations between (i) the expression profile of each gene and the GLS disease scores (reported as gene significance (GS) scores); and (ii) each module eigengene with the GLS disease scores. P-values for the correlations mentioned in (ii) were calculated based on a resampling method (Methods S1). A p-value correction for multiple testing by controlling the false discovery rate (FDR) was conducted for all statistical tests in this paper (Benjamini and Hochberg, 1995).

QTL mapping

QTL mapping for GLS disease severity was conducted using the genetic linkage map QMap 2.0 (Berger *et al.*, 2014) for the CML444 X SC Malawi population. QTL Cartographer (Basten *et al.*, 1994) was used to map QTLs at each of the four ratings (92,

99, 109 and 116 DAP) using the least square mean values of the GLS disease scores. A walking speed of 2 cM was used in composite interval mapping with forward regression and backward elimination (p-value=0.1). Permutation-based LOD score thresholds were calculated per rating to approximate α=0.05 experiment-wise (Churchill and Doerge, 1994).

eQTL data analysis

An eQTL data analysis pipeline (Figure S5; Methods S1) was developed in Python (http://www.python.org) to analyze microarray-based gene expression profiles for 19,281 reporters in the maize leaf samples across 100 RILs of the CML444 X SC Malawi population. It was implemented in the online data analysis platform Galaxy (http://galaxyproject.org) (Giardine *et al.*, 2005), which provides a user-friendly web-based interface for command-line tools. The pipeline, available as three consecutive workflows via the Toolshed (http://toolshed.g2.bx.psu.edu), (i) determines the likelihood ratio (LR) threshold that corrects for genome-wide markers using permutation tests (Churchill and Doerge, 1994) via QTL Cartographer's Zmapqtl module (Basten *et al.*, 1994) and multiple e-traits (in this study, the 95th percentiles for 105 randomly chosen e-traits were determined by 1000 permutations each, and the average LR of the 95th percentiles was taken as the "estimated experiment-wise threshold" (conversion LOD = 0.217 X LR)(Table S16)), (ii) maps eQTLs, using QTL Cartographer (Basten *et al.*, 1994) (with parameters: walking speed of 2 cM; composite interval mapping; forward regression and backward elimination (p-value=0.1); estimated experiment-wise LOD threshold = 2.8), as independent parallel tasks using different nodes on a compute cluster; and (iii) classifies eQTLs as *cis* or *trans* (in this study eQTLs closer than 6.25 cM – half the average size of an eQTL – to its linked gene

were called *cis*-eQTLs) in order to identify significant *trans*-eQTL hotspots (by firstly calculating the genome wide frequency of eQTLs and then normalizing for local gene density). R (R Core Team, 2014) was employed for statistical analysis and data visualisation (mainly in workflow iii).

Meta analysis

Fisher's exact tests with correction for FDR (Benjamini and Hochberg, 1995) were used to determine whether each co-expression module was enriched for genes with eQTLs in each *trans*-eQTL hotspot using a customized script in R (Methods S1, Methods S3).

Module eigengene eQTLs were mapped using parameters described for the global eQTL analysis. A Circos diagram (Krzywinski *et al.*, 2009) was constructed to visualise the integration of QTL, eQTL and co-expression data.

Annotation and GO enrichment analysis

The best *Arabidopsis* TAIR10 and maize annotations were retrieved from the *Zea mays* annotation file, which was released as part of Phytozome version 7.0 (http://www.phytozome.net). Additional annotation information was obtained using Blast2GO and MapMan, which were used to manually classify genes into functional categories (Thimm *et al.*, 2004). BiNGO (Maere *et al.*, 2005) was used to identify enriched GO-terms (http://www.geneontology.org).

RNAseq, RT-qPCR and kauralexin analysis of maize inbred line B73

RNAseq was conducted on leaf samples collected from B73 plants grown in the field at Hildesheim Research Station, South Africa. Three biological replicates each of *C. zeina* infected leaves and control leaves were subjected to RNAseq separately using Illumina technology. Reads were mapped to v5b.60 of the *Z. mays* genome, transcripts were quantified using HTseq (Anders *et al.*, 2015), and differential expression determined with edgeR software (Methods S1 and Methods S4) (Robinson *et al.*, 2010). RNAseq data have been deposited in the NCBI Gene Expression Omnibus (Accession # GSE81344). Fisher's exact tests with correction for FDR were used to determine whether the co-expression modules were enriched for significantly induced or repressed genes in the B73 RNAseq experiment (Methods S1 and Methods S3).

RT-qPCR and kauralexin analysis was conducted on maize B73 plants grown in a glasshouse. At the V8-V10 stage the plants were leaf inoculated with *C. zeina* CMW25467 using methods described in (Korsman *et al.*, 2012). Three biological replicate leaf samples from independent plants were sampled at 0dpi and 18 dpi (after GLS lesions had developed), and subjected to RNA extraction and kauralexin quantification. Kauralexin extraction, derivitisation and GC-MS analysis was carried out according to (Schmelz *et al.*, 2011). RT-qPCR analysis of putative kauralexin biosynthesis genes (GRMZM2G044481; AC214360.3_FGP001 and GRMZM2G161472) is described in Methods S1.

ACKNOWLEDGMENTS

The financial assistance of the Technology Innovation Agency (TIA), the National Research Foundation (NRF) and the University of Pretoria (UP), South Africa, towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at,

are those of the authors and are not necessarily to be attributed to TIA, NRF or UP. We thank P Tongoona and F Middleton for assistance with the maize field trials and GLS disease severity analysis, JN Korsman and E Wentzel for assistance with plant molecular biology experiments, and two anonymous reviewers for their insights. The authors d clare no conflicts of interest.

SUPPORTING INFORMATION

Figure S1. Systems genetics strategy.

Figure S2. Graph of eigengene values for *GY-s* and *TQ-r* gene co-expression modules in the maize RIL population.

Figure S3. Genomic distribution of eQTL from the maize RIL population.

Figure S4. *Trans*-eQTL distribution and hot-spots in the maize RIL population.

Figure S5. Flow diagram of the eQTL data analysis pipeline.

Table S1. Co-expression module membership of all 19,281 reporters and correlation of reporter expression with GLS disease scores across 100 RILs of the CML444 X SC Malawi population.

Table S2. Co-expression modules identified by WGCNA of microarray data from 100 RILs of the CML444 X SC Malawi population.

Table S3. Reporters in the *GY-s* co-expression module of the maize RIL population, and expression in B73-*C. zeina*.

Table S4. Reporters and enriched GO term in the *PT-s* co-expression module of the maize RIL population.

Table S5. Reporters in the *TQ-r* co-expression module of the maize RIL population, and expression in B73-*C. zeina*.

Table S6. Enriched GO-terms for the *TQ-r* module.

Table S7. All eQTLs identified from microarray data from the CML444 X SC Malawi maize population.

Table S8. Global eQTL summary and *cis*/*trans*-eQTL classification.

Table S9. All *trans*-eQTL hotspots, and co-expression modules enriched for reporters in these *trans*-eQTL hotspots.

Table S10. Reporters with *cis*-eQTLs that overlap the GLS QTLs.

Table S11. Reporters in the *GY-s* module with eQTLs in eQTL_HS9a(S) and/or eQTL_HS10c(S).

Table S12. Reporters in the *TQ-r* module with eQTLs in eQTL_HS9a(R) and eQTL_HS9b(R).

Table S13. Differentially expressed genes in maize inbred B73 challenged with *C. zeina* (RNAseq data).

Table S14.1 GO enrichment of B73 genes significantly induced by *C. zeina* (RNAseq; FDR<0.05) that are present in the *GY-s* co-expression module.

Table S14.2 Maize B73 genes significantly induced by *C. zeina* (RNAseq; FDR<0.05) that are present in the *GY-s* co-expression module.

Table S15. Comparison of co-expression module gene lists and list of DEGs in B73 challenged with *C. zeina.*

Table S16. eQTL permutation workflow output for 105 randomly selected expression traits.

Methods S1. Additional experimental procedures.

Methods S2. R-script for WGCNA of maize RIL population microarray data.

Methods S3. R-script for Fisher's Exact tests to determine whether co-expression modules were enriched for genes with eQTLs in *trans*-eQTL hotspots.

Methods S4. R-script for EdgeR analysis of maize B73 RNAseq data.

REFERENCES

Anders, S., Pyl, P.T. and Huber, W. (2015) HTSeq – A Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.

Basten, C.J., Weir, B.S. and Zeng, Z.-B. (1994) Zmap-a QTL cartographer. In C. Smith, J. S. Gavora, B. Benkel, J. Chesnais, W. Fairfull, J. P. Gibson, B. W. Kennedy, and E. B. Burnside, eds. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software*. the Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada, pp. 65–66.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.*, **57**, 289–300.

Benson, J.M., Poland, J.A., Benson, B.M., Stromberg, E.L. and Nelson, R.J. (2015) Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. *PLOS Genet.*, **11**, e1005045.

Berger, D.K., Carstens, M., Korsman, J.N., Middleton, F., Kloppers, F.J., Tongoona, P. and Myburg, A.A. (2014) Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*. *BMC Genet.*, **15**, 60.

Chen, X., Hackett, C.A., Niks, R.E., et al. (2010) An eQTL analysis of partial resistance to *Puccinia hordei* in barley. *PLoS One*, **5**, e8598.

Churchill, G.A. and Doerge, R.W. (1994) Empirical Threshold Values for Quantitative Trait Mapping. *Genetics*, **138**, 963–971.

Civelek, M. and Lusis, A.J. (2014) Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.*, **15**, 34–48.

Claverie, M., Souquet, M., Jean, J., Forestier-Chiron, N., Lepitre, V., Pré, M., Jacobs, J., Llewellyn, D. and Lacape, J.-M. (2012) cDNA-AFLP-based genetical genomics in cotton fibers. *Theor. Appl. Genet.*, **124**, 665–683.

Coetzer, N., Myburg, A.A. and Berger, D.K. (2011) Maize microarray annotation database. *Plant Methods*, **7**, 31.

Desjardins, A., Gardner, H. and Weltring, K. (1992) Detoxification of sesquiterpene phytoalexins by Gibberella pulicaris (*Fusarium sambucinum*) and its importance for

virulence on potato tubers. *J. Ind. Microbiol.*, **9**, 201–211.

Devaiah, B.N., Karthikeyan, A.S. and Raghothama, K.G. (2007) WRKY75 transcription factor is a modulator of phosphate acquisition and root development in Arabidopsis. *Plant Physiol.*, **143**, 1789–1801.

Drost, D.R., Puranik, S., Novaes, E., Novaes, C.R.D.B., Dervinis, C., Gailing, O. and Kirst, M. (2015) Genetical genomics of Populus leaf shape variation. *BMC Plant Biol.*, **15**, 166.

Druka, A., Potokina, E., Luo, Z., et al. (2008) Exploiting regulatory variation to identify genes underlying quantitative resistance to the wheat stem rust pathogen *Puccinia graminis f. sp. tritici* in barley. *Theor. Appl. Genet.*, **117**, 261–272.

Ellinger, D., Glöckner, A., Koch, J., Naumann, M., Stürtz, V., Schütt, K., Manisseri, C., Somerville, S.C. and Voigt, C.A. (2014) Interaction of the Arabidopsis GTPase RabA4c with Its Effector PMR4 Results in Complete Penetration Resistance to Powdery Mildew. *Plant Cell*, **26**, 3185–3200.

Ellinger, D., Naumann, M., Falter, C., Zwikowics, C., Jamrow, T., Manisseri, C., Somerville, S.C. and Voigt, C.A. (2013) Elevated early callose deposition results in complete penetration resistance to powdery mildew in Arabidopsis. *Plant Physiol.*, **161**, 1433–1444.

Feltus, F.A. (2014) Systems genetics: A paradigm to improve discovery of candidate genes and mechanisms underlying complex traits. *Plant Sci.*, **223**, 45–48.

Furniss, J.J. and Spoel, S.H. (2015) Cullin-RING ubiquitin ligases in salicylic acid-mediated plant immune signaling. *Front. Plant Sci.*, **6**, 1–10.

Giardine, B., Riemer, C., Hardison, R.C., et al. (2005) Galaxy: A platform for interactive large-scale genome analysis. *Genome Res.*, **15**, 1451–1455.

Glazebrook, J. (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu. Rev. Phytopathol.*, **43**, 205–227. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16078883 [Accessed March 6, 2012].

Hansen, B.G., Halkier, B.A. and Kliebenstein, D.J. (2008) Identifying the molecular basis of QTLs: eQTLs add a new dimension. *Trends Plant Sci.*, **13**, 72–77.

He, Y., Chung, E.-H., Hubert, D.A., Tornero, P. and Dangl, J.L. (2012) Specific missense alleles of the Arabidopsis jasmonic acid co-receptor COI1 regulate innate immune receptor accumulation and function. *PLoS Genet.*, **8**, e1003018.

Hinch, J. and Clarke, A.E. (1982) Callose formation in *Zea mays* as a response to infection with *Phytophthora cinnamomi*. *Physiol. Plant Pathol.*, **21**, 113–124.

Holloway, B., Luck, S., Beatty, M., Rafalski, J.-A. and Li, B. (2011) Genome-wide expression quantitative trait loci (eQTL) analysis in maize. *BMC Genomics*, **12**, 336.

Hsu, K.H., Liu, C.C., Wu, S.J., et al. (2014) Expression of a gene encoding a rice RING zinc-finger protein, OsRZFP34, enhances stomata opening. *Plant Mol. Biol.*, **86**, 125–137.

Hurni, S., Scheuermann, D., Krattinger, S.G., et al. (2015) The maize disease resistance gene Htn1 against northern corn leaf blight encodes a wall-associated receptor-like kinase. *Proc. Natl. Acad. Sci.*, **112**, 8780–8785.

Jamann, T., Nelson, R. and Balint-kurti, P. (2013) The genetic basis of disease resistance in maize. In R. K. Varshney and R. Tuberosa, eds. *Translational genomics for crop breeding: biotic stress.* John Wiley & Sons, Inc. Published, pp. 31–43.

Jansen, R.C. and Nap, J.P. (2001) Genetical genomics: the added value from segregation. *Trends Genet.*, **17**, 388–391.

Jiang, S., Yao, J., Ma, K.-W., Zhou, H., Song, J., He, S.Y. and Ma, W. (2013) Bacterial Effector Activates Jasmonate Signaling by Directly Targeting JAZ Transcriptional Repressors. *PLoS Pathog.*, **9**, e1003715.

Jiménez-Gómez, J.M., Wallace, A.D. and Maloof, J.N. (2010) Network analysis identifies *ELF3* as a QTL for the shade avoidance response in Arabidopsis. *PLoS Genet.*, **6**, e1001100.

Jines, M.P., Balint-Kurti, P., Robertson-Hoyt, L.A., Molnar, T., Holland, J.B. and Goodman, M.M. (2007) Mapping resistance to Southern rust in a tropical by temperate maize recombinant inbred topcross population. *Theor. Appl. Genet.*, **114**, 659–667.

Jordan, M.C., Somers, D.J. and Banks, T.W. (2007) Identifying regions of the wheat genome controlling seed development by mapping expression quantitative trait loci. *Plant Biotechnol. J.*, **5**, 442–53.

Keurentjes, J.J.B., Fu, J., Terpstra, I.R., et al. (2007) Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proc. Natl. Acad. Sci.*, **104**, 1708–1713.

Kim, H., Ridenour, J.B., Dunkle, L.D. and Bluhm, B.H. (2011) Regulation of pathogenesis by light in *Cercospora zeae-maydis*: an updated perspective. *Plant Pathol. J.*, **27**, 103–109.

Kirst, M., Basten, C.J., Myburg, A.A., Zeng, Z.-B. and Sederoff, R.R. (2005) Genetic architecture of transcript-level variation in differentiating xylem of a eucalyptus hybrid. *Genetics*, **169**, 2295–303.

Kloosterman, B., Kumari, A.M., Chibon, P.-Y., Oortwijn, M., Linden, G.C. van der, Visser,

R.G.F. and Bachem, C.W.B. (2012) Organ specificity and transcriptional control of metabolic routes revealed by expression QTL profiling of source--sink tissues in a segregating potato population. *BMC Plant Biol.*, **12**, 17.

Korasick, D.A., McMichael, C., Walker, K.A., Anderson, J.C., Bednarek, S.Y. and Heese, A. (2010) Novel Functions of Stomatal Cytokinesis-Defective 1 (SCD1) in Innate Immune Responses against Bacteria. *J. Biol. Chem.*, **285**, 23342–23350.

Korsman, J., Meisel, B., Kloppers, F.J., Crampton, B.G. and Berger, D.K. (2012) Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR. *Eur. J. Plant Pathol.*, **133**, 461–471.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J. and Marra, M.A. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.*, **19**, 1639–1645.

Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.

Li, L., Petsch, K., Shimizu, R., et al. (2013) Mendelian and non-Mendelian regulation of gene expression in maize. *PLoS Genet.*, **9**, e1003202.

Li, Y., Pearl, S.A. and Jackson, S.A. (2015) Gene networks in plant biology: approaches in reconstruction and analysis. *Trends Plant Sci.*, **20**, 664–675.

Lingner, U., Münch, S., Deising, H.B. and Sauer, N. (2011) Hexose Transporters of a Hemibiotrophic Plant Pathogen: FUNCTIONAL VARIATIONS AND REGULATORY DIFFERENCES AT DIFFERENT STAGES OF INFECTION. *J. Biol. Chem.*, **286**, 20913–20922.

Liu, H., Sun, M., Du, D., Pan, H., Cheng, T., Wang, J., Zhang, Q. and Gao, Y. (2016) Whole-transcriptome analysis of differentially expressed genes in the vegetative buds, floral buds and buds of *Chrysanthemum morifolium*. *BMC Genomics*, **17**.

Liu, K.-J. and Xu, X.-D. (2013) First report of gray leaf spot of maize caused by *Cercospora zeina* in China. *Plant Dis.*, **97**, 1656.

Maere, S., Heymans, K. and Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448–3449.

Mammadov, J., Sun, X., Gao, Y., et al. (2015) Combining powers of linkage and association mapping for precise dissection of QTL controlling resistance to gray leaf spot disease in maize (*Zea mays* L.). *BMC Genomics*, **16**, 916.

Meisel, B., Korsman, J., Kloppers, F.J. and Berger, D.K. (2009) *Cercospora zeina* is the causal agent of grey leaf spot disease of maize in southern Africa. *Eur. J. Plant Pathol.*, **124**, 577–583.

Messmer, R., Fracheboud, Y., Bänziger, M., Vargas, M., Stamp, P. and Ribaut, J.-M. (2009) Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. *Theor. Appl. Genet.*, **119**, 913–930.

Moscou, M.J., Lauter, N., Steffenson, B. and Wise, R.P. (2011) Quantitative and qualitative stem rust resistance factors in barley are associated with transcriptional suppression of defense regulons. *PLoS Genet.*, **7**, e1002208.

Munkvold, J.D., Laudencia-Chingcuanco, D. and Sorrells, M.E. (2013) Systems genetics of environmental response in the mature wheat embryo. *Genetics*, **194**, 265–77.

Murray, S.L., Ingle, R. a, Petersen, L.N. and Denby, K.J. (2007) Basal resistance against Pseudomonas syringae in Arabidopsis involves WRKY53 and a protein with homology to a nematode resistance protein. *Mol. Plant-Microbe Interact.*, **20**, 1431–1438.

Nagels Durand, A., Pauwels, L. and Goossens, A. (2016) The Ubiquitin System and Jasmonate Signaling. *Plants*, **5**, 6.

Neves, D.L., Silva, C.N., Pereira, C.B., Campos, H.D. and Tessmann, D.J. (2015) *Cercospora zeina* is the main species causing gray leaf spot in southern and central Brazilian maize regions. *Trop. Plant Pathol.*, **40**, 368–374.

Piisilä, M., Keceli, M.A., Brader, G., Jakobson, L., Jõesaar, I., Sipari, N., Kollist, H., Palva, E.T. and Kariola, T. (2015) The F-box protein MAX2 contributes to resistance to bacterial phytopathogens in *Arabidopsis thaliana*. *BMC Plant Biol.*, **15**, 1–17.

Poland, J.A., Balint-Kurti, P.J., Wisser, R.J., Pratt, R.C. and Nelson, R.J. (2009) Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci.*, **14**, 21–29.

Potokina, E., Druka, A., Luo, Z., Wise, R., Waugh, R. and Kearsey, M. (2008) Gene expression quantitative trait locus analysis of 16 000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. *Plant J.*, **53**, 90–101.

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

Rudd, J.J., Kanyuka, K., Hassani-Pak, K., et al. (2015) Transcriptome and metabolite profiling the infection cycle of Zymoseptoria tritici on wheat (Triticum aestivum) reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions, and a variation on the hemibiotro. *Plant Physiol.*, **167**, 1158–1185.

Schadt, E.E., Monks, S.A., Drake, T.A., et al. (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature*, **422**, 279–302.

Schirawski, J. (2015) Invasion is sweet. *New Phytol.*, **206**, 892–894.

Schmelz, E.A., Kaplan, F., Huffaker, A., Dafoe, N.J., Vaughan, M.M., Ni, X., Rocca, J.R., Alborn, H.T. and Teal, P.E. (2011) Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proc. Natl. Acad. Sci.*, **108**, 5455–5460.

Schweiger, W., Steiner, B., Ametz, C., et al. (2013) Transcriptomic characterization of two major *Fusarium* resistance quantitative trait loci (QTLs), *Fhb1* and *Qfhs.ifa-5A*, identifies novel candidate genes. *Mol. Plant Pathol.*, **14**, 772–785.

Seybold, H., Trempel, F., Ranf, S., Scheel, D., Romeis, T. and Lee, J. (2014) Ca2+ signalling in plant immune response: From pattern recognition receptors to Ca2+ decoding mechanisms. *New Phytol.*, **204**, 782–790.

Shi, C., Uzarowska, A., Ouzunova, M., Landbeck, M., Wenzel, G. and Lübberstedt, T. (2007) Identification of candidate genes associated with cell wall digestibility and eQTL (expression quantitative trait loci) analysis in a Flint x Flint maize recombinant inbred line population. *BMC Genomics*, **8**, 22.

Shi, L., Lv, X., Weng, J., et al. (2014) Genetic characterization and linkage disequilibrium mapping of resistance to gray leaf spot in maize (*Zea mays* L.). *Crop J.*, 1–12.

Shim, W.B. and Dunkle, L.D. (2002) Identification of genes expressed during cercosporin biosynthesis in *Cercospora zeae-maydis*. *Physiol. Mol. Plant Pathol.*, **61**, 237–248.

St.Clair, D.A. (2010) Quantitative disease resistance and quantitative resistance loci in breeding. *Annu. Rev. Phytopathol.*, **48**, 247–268.

Swanson-Wagner, R.A., DeCook, R., Jia, Y., Bancroft, T., Ji, T., Zhao, X., Nettleton, D. and Schnable, P.S. (2009) Paternal dominance of trans-eQTL influences gene expression patterns in maize hybrids. *Science (80-.).*, **326**, 1118–1120.

Thimm, O., Bläsing, O., Gibon, Y., et al. (2004) Mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.*, **37**, 914–939.

Thines, B., Katsir, L., Melotto, M., et al. (2007) JAZ repressor proteins are targets of the SCFCOI1 complex during jasmonate signalling. *Nature*, **448**, 661–665.

Trachsel, S., Messmer, R., Stamp, P. and Hund, A. (2009) Mapping of QTLs for lateral and axile root growth of tropical maize. *Theor. Appl. Genet.*, **119**, 1413–1424.

Trachsel, S., Messmer, R., Stamp, P., Ruta, N. and Hund, A. (2010) QTLs for early vigor of tropical maize. *Mol. Breed.*, **25**, 91–103.

Tsuda, K. and Somssich, I.E. (2015) Transcriptional networks in plant immunity. *New Phytol.*, **206**, 932–947.

Vanderplank, J. (1984) *Diease resistance in plants.* Second Edi., London: Academic Press, Inc.

Vargas, W.A., Martín, J.M.S., Rech, G.E., Rivera, L.P., Benito, E.P., Díaz-Mínguez, J.M., Thon, M.R. and Sukno, S.A. (2012) Plant defense mechanisms are activated during biotrophic and necrotrophic development of *Colletotricum graminicola* in maize. *Plant Physiol.*, **158**, 1342–1358.

Wang, J., Levy, M. and Dunkle, L.D. (1998) Sibling species of *Cercospora* associated with gray leaf spot of maize. *Phytopathology*, **88**, 1269–1275.

Wang, J., Yu, H., Weng, X., Xie, W., Xu, C., Li, X., Xiao, J. and Zhang, Q. (2014) An expression quantitative trait loci-guided co-expression analysis for constructing regulatory network using a rice recombinant inbred line population. *J. Exp. Bot.*, **65**, 1069–1079.

Wang, J., Yu, H., Xie, W., Xing, Y., Yu, S., Xu, C., Li, X., Xiao, J. and Zhang, Q. (2010) A global analysis of QTLs for expression variations in rice shoots at the early seedling stage. *Plant J.*, **63**, 1063–1074.

Ward, J.M.J., Stromberg, E.L., Nowell, D.C. and Nutter, F.W. (1999) Gray leaf spot: a disease of global importance in maize production. *Plant Dis.*, **83**, 884–895.

Wentzell, A.M., Rowe, H.C., Hansen, B.G., Ticconi, C., Halkier, B.A. and Kliebenstein, D.J. (2007) Linking Metabolic QTLs with Network and cis-eQTLs Controlling Biosynthetic Pathways. *PLoS Genet.*, **3**, 1687–1701.

West, M.A.L., Kim, K., Kliebenstein, D.J., Leeuwen, H. van, Michelmore, R.W., Doerge, R.W. and St Clair, D.A. (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. *Genetics*, **175**, 1441–1450.

Wisser, R.J., Balint-Kurti, P.J. and Nelson, R.J. (2006) The genetic architecture of disease resistance in maize: a synthesis of published studies. *Phytopathology*, **96**, 120–129.

Wisser, R.J., Kolkman, J.M., Patzoldt, M.E., Holland, J.B., Yu, J., Krakowsky, M., Nelson, R.J.

and Balint-Kurti, P.J. (2011) Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a GST gene. *Proc. Natl. Acad. Sci.*, **108**, 7339–7344.

Xu, L., Zhang, Y., Shao, S., Chen, W., Tan, J., Zhu, M., Zhong, T., Fan, X. and Xu, M. (2014) High-resolution mapping and characterization of qRgls2, a major quantitative trait locus involved in maize resistance to gray leaf spot. *BMC Plant Biol.*, **14**, 230.

Zhang, B. and Horvath, S. (2005) A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.*, **4**, Article17.

Zhang, F., Yao, J., Ke, J., et al. (2015) Structural basis of JAZ repression of MYC transcription factors in jasmonate signalling. *Nature*, **525**, 269–273.

Appendix S1. Legends for Supporting Material

Figure S1 Systems genetics strategy.

(A) Gene co-expression analysis, i.e. microarray gene expression profiles of leaf material from *C. zeina*-infected plants of the CML444 X SC Malawi maize RIL population were subjected to weighted gene co-expression analysis (WCGNA) to identify modules of co-expressed genes. Phenotypes (GLS disease severity measurements) across the same individuals were incorporated to identify co-expression modules that correlated positively or negatively with GLS severity. (B) Global eQTL analysis with the same microarray expression data was performed. (C) QTL analysis for the phenotype (GLS severity) was carried out. (D) Overlap analysis to identify regions in the genome where GLS severity QTL overlapped with trans eQTL hotspots. (E) Meta-analysis combined co-expression analysis with the QTL/*trans*-eQTL hotspot overlap analysis in order to arrive at a genetic basis for the observed coordinated expression responses to GLS disease.

Figure S2. Graph of eigengene values for *GY-s* and *TQ-r* gene co-expression modules in the maize RIL population. The y-axis scale is the log expression values of the module eigengenes, and the x-axis scale is the GLS disease severity score for the 100 RILs of the CML444 X SC Malawi population. The RILs have been sorted from the most resistant (GLS score = 1) to most susceptible (GLS score = 9)(left to right). The green line represents the fitted regression line through the *GY-s* module eigengene log expression values, with deviations from the line shown by the filled green colour. Reporters in the *GY-s* module are positively correlated with GLS severity across the RILs (correlation = 0.71; FDR-adjusted p-value = 4E-15), i.e. RILs with low GLS disease scores (resistant RILs) have with low module eigengene expression values, and RILs with high GLS disease scores have high module eigengene expression values. Reporters in the *TQ-r* module show the opposite behavior (correlation = -0.31; FDR-adjusted

p-value = 0.03), with a fitted regression line through the *TQ-r* module eigengene log expression values shown by a blue line, and deviations from the line shown by the filled blue colour.

Figure S3 Genomic distribution of eQTL from the maize RIL population. Scatter plot of the genomic relationships between eQTL positions (x-axis) and the corresponding expression trait (reporter) positions (y-axis) on the maize genome for the CML444 X SC Malawi maize RIL population grown in the field with GLS disease pressure. The figure was generated by the "classification" component of the *Backend workflow* of the eQTL data analysis pipeline. The color-key distinguishes between *cis*-eQTL (blue) and *trans*-eQTL (green). The ten maize chromosomes are separated by grey dashed lines.

Figure S4. *Trans*-eQTL distribution and hot-spots in the maize RIL population. Distribution of *trans*-eQTLs per cM and identification of *trans*-eQTL hotspots across the ten maize chromosomes for the CML444 X SC Malawi maize RIL population grown in the field with GLS disease pressure. This figure was generated by the "hotspots" component of the *Backend workflow*. The y-axis gives the number of *trans*-eQTLs per cM that was calculated per sliding window bin (x-axis). The horizontal line shows the permuted threshold (p-value < 0.05) for detection of *trans*-eQTL hotspots. Sliding window bins where the number of eQTLs per cM was above the permutation threshold and with significant eQTL excess compared to gene number (chi-squared test p-value< 0.0001), were declared as "unbiased" eQTL hotspots and marked in red.

Figure S5. Flow diagram of the eQTL data analysis pipeline implemented in Galaxy, described in Methods S1.

Table S1. Co-expression module membership of all 19,281 reporters and correlation of reporter expression with GLS disease scores across 100 RILs of the CML444 X SC Malawi population.

Table S2. Co-expression modules identifed by WGCNA of microarray data from 100 RILs of the CML444 X SC Malawi population.

Table S3. Reporters in the *GY-s* co-expression module of the maize RIL population, and expression in B73-*C. zeina*.

Table S4. Reporters and enriched GO term in the *PT-s* co-expression module of the maize RIL population.

Table S5. Reporters in the *TQ-r* co-expression module of the maize RIL population, and expression in B73-*C. zeina*.

Table S6. Enriched GO-terms for the *TQ-r* module.

Table S7. All eQTLs identified from microarray data from the CML444 X SC Malawi maize population.

Table S8. Global eQTL summary and cis/trans-eQTL classification.

Table S9. All *trans*-eQTL hotspots, and co-expression modules enriched for reporters in these *trans*-eQTL hotspots.

Table S10. Reporters with cis-eQTLs that overlap the GLS QTLs.

Table S11. Reporters in the *GY-s* module with eQTLs in eQTL_HS9a(S) and/or eQTL_HS10c(S).

Table S12. Reporters in the *TQ-r* module with eQTLs in eQTL_HS9a(R) and eQTL_HS9b(R).

Table S13. Differentially expressed genes in maize inbred B73 challenged with *C. zeina* (RNAseq data).

Table S14.1 GO enrichment of B73 genes significantly induced by *C. zeina* (RNAseq; FDR<0.05) that are present in the *GY-s* co-expression module

Table S14.2 Maize B73 genes significantly induced by *C. zeina* (RNAseq; FDR<0.05) that are present in the *GY-s* co-expression module

Table S15. Comparison of co-expression module gene lists and list of DEGs in B73 challenged with *C. zeina.*

Table S16. eQTL permutation workflow output for 105 randomly selected expression traits.

Methods S1. Additional experimental procedures.

Methods S2. R-script for WGCNA of maize RIL population microarray data.

Methods S3. R-script for Fisher's Exact tests to determine whether co-expression modules were enriched for genes with eQTLs in *trans*-eQTL hotspots.

Methods S4. R-script for EdgeR analysis of maize B73 RNAseq data.

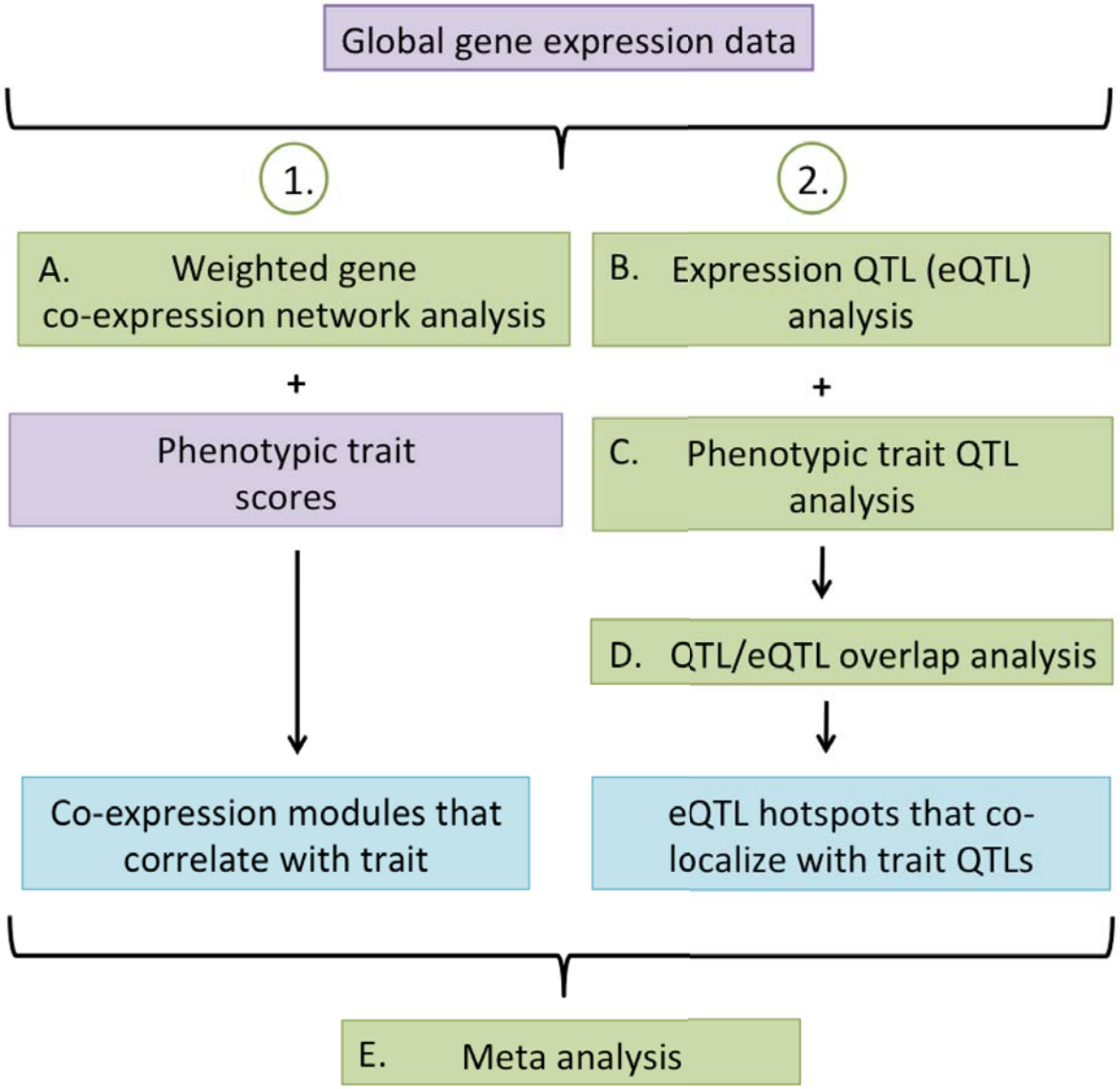

Figure S1 Systems genetics strategy.
(A) Gene co-expression analysis, i.e. microarray gene expression profiles of leaf material from *C. zeina*-infected plants of the CML444 X SC Malawi maize RIL population were subjected to weighted gene co-expression analysis (WCGNA) to identify modules of co-expressed genes. Phenotypes (GLS disease severity measurements) across the same individuals were incorporated to identify co-expression modules that correlated positively or negatively with GLS severity. (B) Global eQTL analysis with the same microarray expression data was performed. (C) QTL analysis for the phenotype (GLS severity) was carried out. (D) Overlap analysis to identify regions in the genome where GLS severity QTL overlapped with trans eQTL hotspots. (E) Meta-analysis combined co-expression analysis with the QTL/trans-eQTL hotspot overlap analysis in order to arrive at a genetic basis for the observed coordinated expression responses to GLS disease.

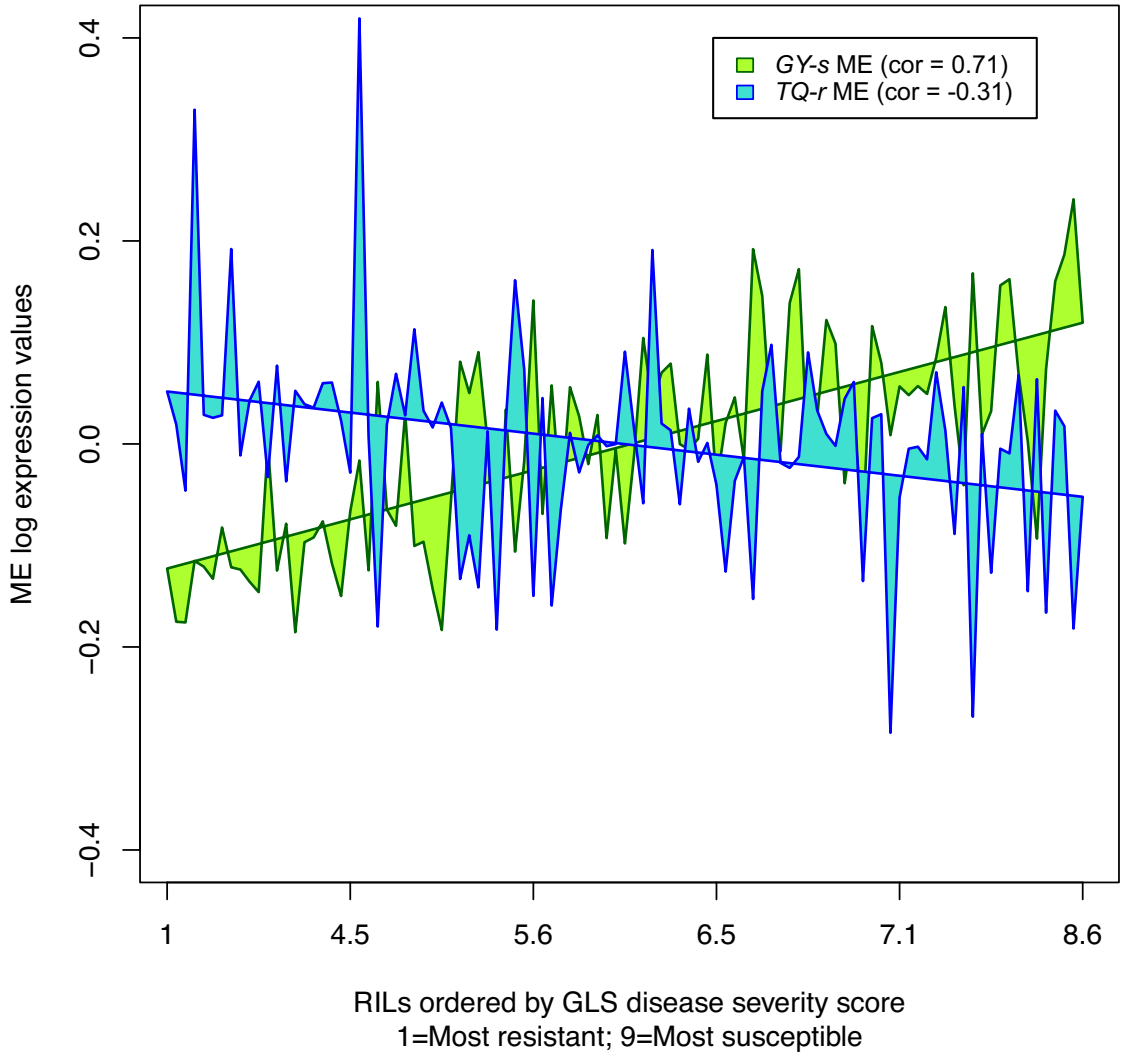

Figure S2 Graph of eigengene values for *GY-s* and *TQ-r* gene co-expression modules in the maize RIL population. The y-axis scale is the log expression values of the module eigengenes, and the x-axis scale is the GLS disease severity score for the 100 RILs of the CML444 X SC Malawi population. The RILs have been sorted from the most resistant (GLS score = 1) to most susceptible (GLS score = 9)(left to right). The green line represents the fitted regression line through the *GY-s* module eigengene log expression values, with deviations from the line shown by the filled green colour. Reporters in the *GY-s* module are positively correlated with GLS severity across the RILs (correlation = 0.71; FDR-adjusted p-value = 4E-15), i.e. RILs with low GLS disease scores (resistant RILs) have with low module eigengene expression values, and RILs with high GLS disease scores have high module eigengene expression values. Reporters in the *TQ-r* module show the opposite behavior (correlation = -0.31; FDR-adjusted p-value = 0.03), with a fitted regression line through the *TQ-r* module eigengene log expression values shown by a blue line, and deviations from the line shown by the filled blue colour.

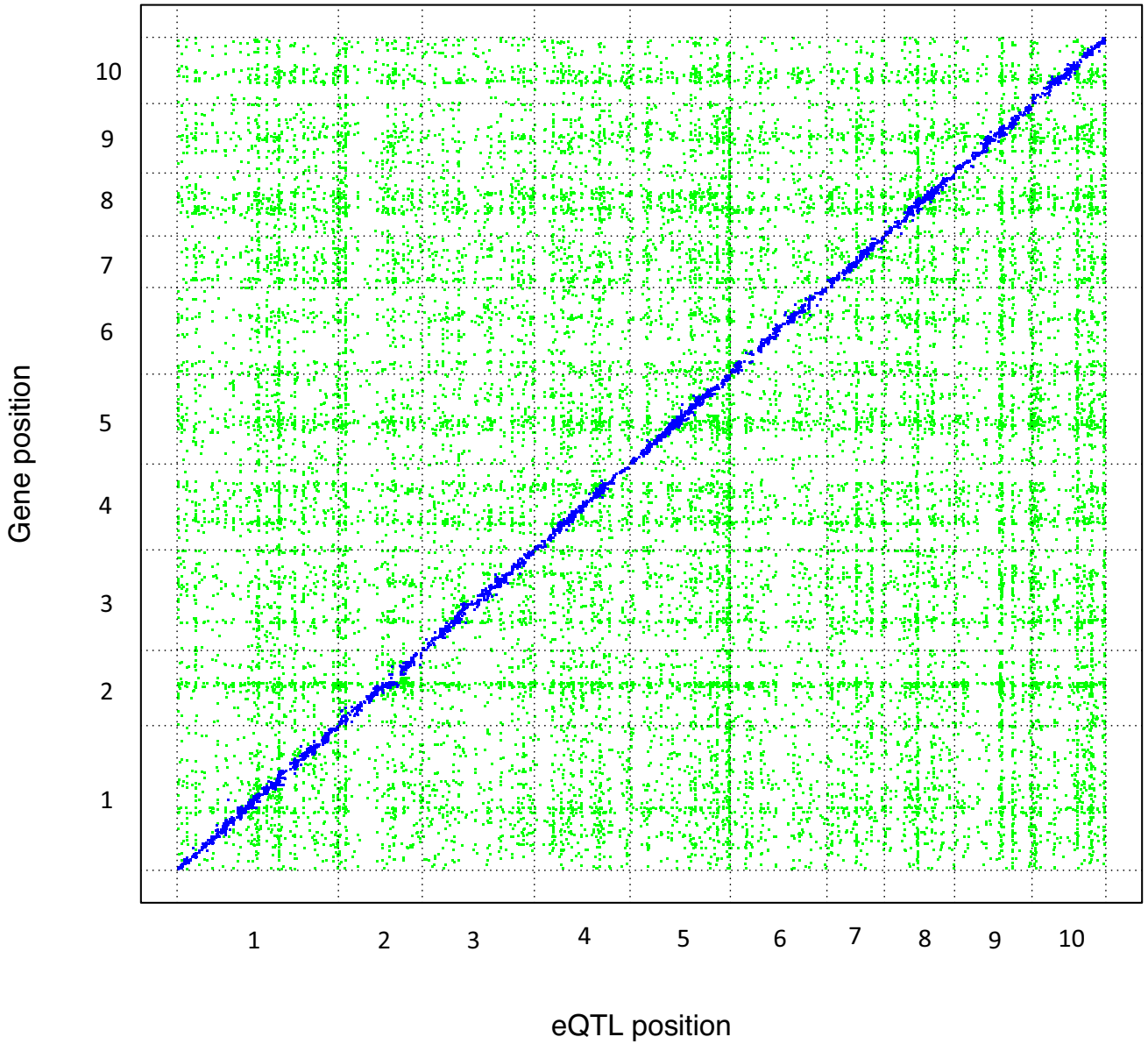

Figure S3 Genomic distribution of eQTL from the maize RIL population. Scatter plot of the genomic relationships between eQTL positions (x-axis) and the corresponding expression trait (reporter) positions (y-axis) on the maize genome for the CML444 X SC Malawi maize RIL population grown in the field with GLS disease pressure. The figure was generated by the "classification" component of the *Backend workflow* of the eQTL data analysis pipeline. The color-key distinguishes between *cis*-eQTL (blue) and *trans*-eQTL (green). The ten maize chromosomes are separated by grey dashed lines.

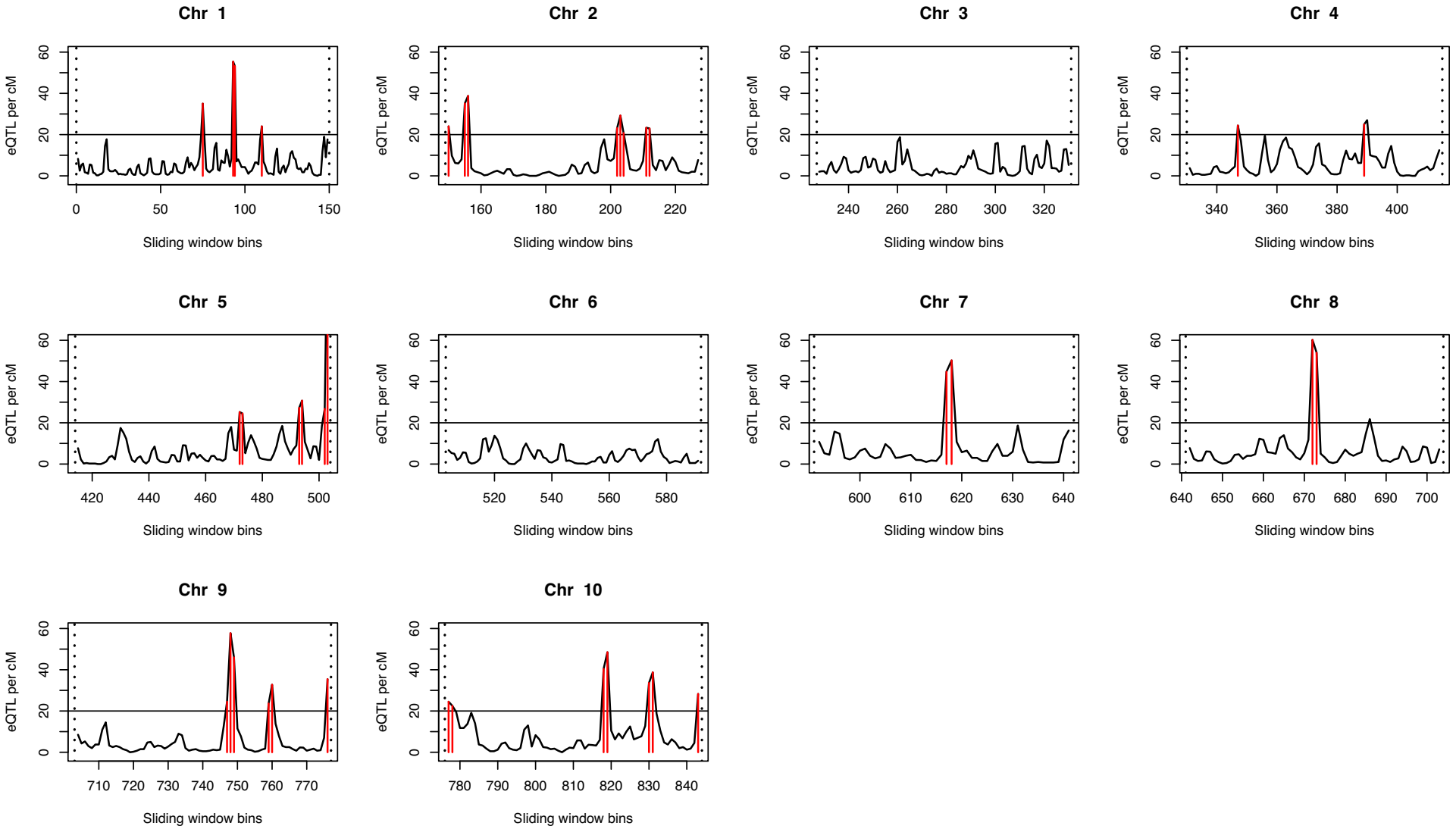

Figure S4 *Trans*-eQTL distribution and hot-spots in the maize RIL population. Distribution of *trans*-eQTLs per cM and identification of *trans*-eQTL hotspots across the ten maize chromosomes for the CML444 X SC Malawi maize RIL population grown in the field with GLS disease pressure. This figure was generated by the "hotspots" component of the *Backend workflow*. The y-axis gives the number of *trans*-eQTLs per cM that was calculated per sliding window bin (x-axis). The horizontal line shows the permuted threshold (p-value < 0.05) for detection of trans-eQTL hotspots. Sliding window bins where the number of eQTLs per cM was above the permutation threshold and with significant eQTL excess compared to gene number (chi-squared test p-value< 0.0001), were declared as "unbiased" eQTL hotspots and marked in red.

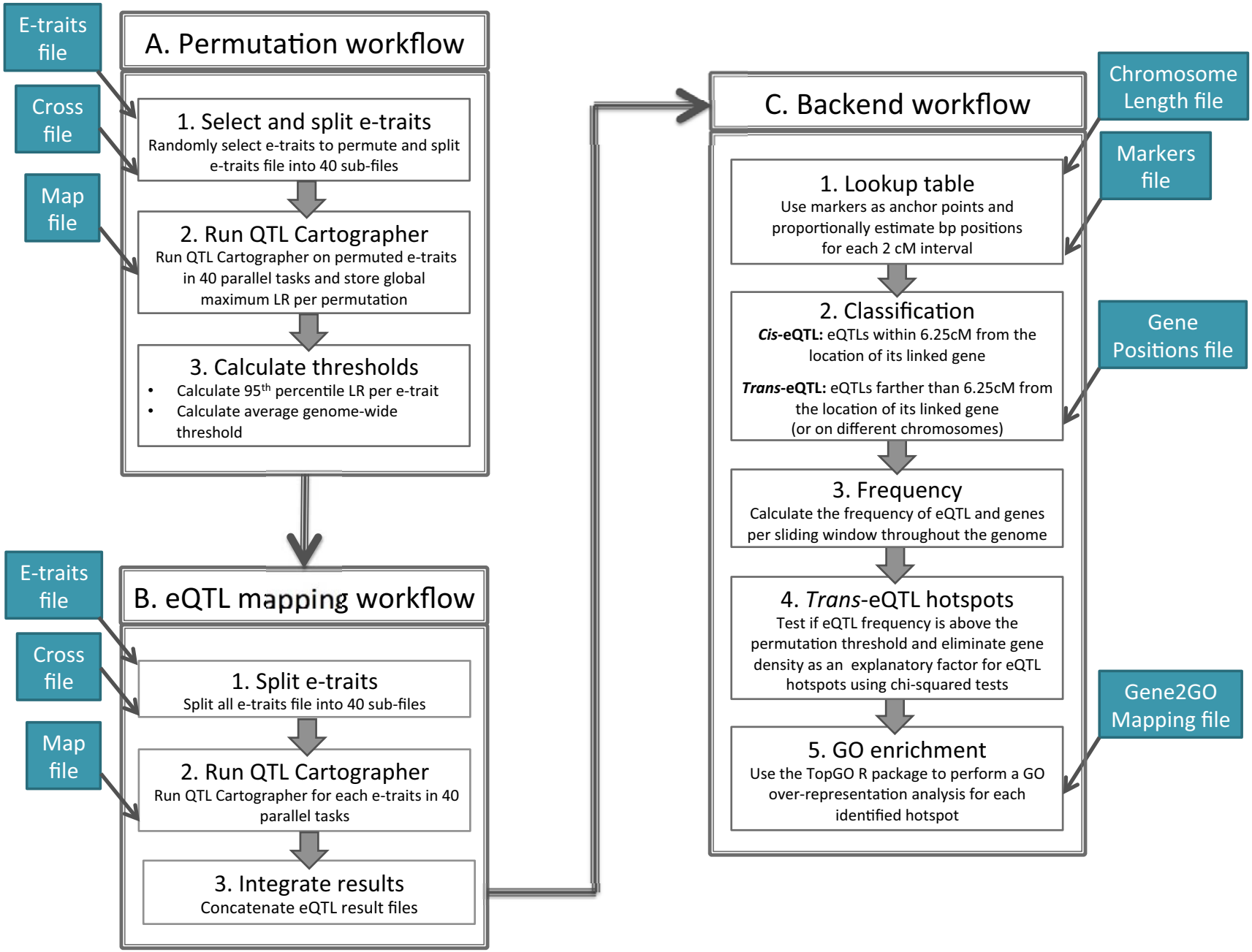

Figure S5 Flow diagram of the eQTL data analysis pipeline implemented in Galaxy, described in Methods S1.