Integration of GWAS and TWAS to elucidate the genetic architecture of natural variation for leaf cuticular conductance in maize

Meng Lin[1], Pengfei Qiao[2], Susanne Matschi[3†], Miguel Vasquez[3], Guillaume P. Ramstein[4§], Richard Bourgault[5], Marc Mohammadi[5], Michael J. Scanlon[2], Isabel Molina[5], Laurie G. Smith[3], and Michael A. Gore[1,*]

[1]Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

[2]Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

[3]Section of Cell and Developmental Biology, University of California San Diego, La Jolla, CA 92093, USA

[4]Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853, USA

[5]Department of Biology, Algoma University, Sault Ste Marie, ON P6A 2G4, Canada

[*]Corresponding authors: Michael A. Gore, 358 Plant Science Building, Cornell University, Ithaca, NY 14853, USA E-mail: mag87@cornell.edu

[†]Present address: Leibniz Institute of Plant Biochemistry, 06120 Halle (Saale), Germany

[§]Present address: Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus Denmark, 8000

Short title

GWAS/TWAS of maize leaf cuticular conductance

Sentence summary

We performed an integrated GWAS/TWAS and identified 22 candidate genes putatively regulating variation in maize leaf g_c. The association between g_c and the strongest candidate causal gene, *ISTL1*, was validated with maize nested near-isogenic lines.

Keywords

Leaf cuticular conductance, Genome-wide association study, Transcriptome-wide association study, near-isogenic lines

Abstract

The cuticle, a hydrophobic layer of cutin and waxes synthesized by plant epidermal cells, is the major barrier to water loss when stomata are closed. Dissecting the genetic architecture of natural variation for maize leaf cuticular conductance (g_c) is important for identifying genes relevant to improving crop productivity in drought-prone environments. To this end, we performed an integrated genome- and transcriptome-wide association study (GWAS/TWAS) to identify candidate genes putatively regulating variation in leaf g_c. Of the 22 plausible candidate genes identified, five were predicted to be involved in cuticle precursor biosynthesis and export, two in cell wall modification, nine in intracellular membrane trafficking, and seven in the regulation of cuticle development. A gene encoding an INCREASED SALT TOLERANCE1-LIKE1 (ISTL1) protein putatively involved in intracellular protein and membrane trafficking was identified in GWAS and TWAS as the strongest candidate causal gene. A set of maize nested near-isogenic lines that harbor the *ISTL1* genomic region from eight donor parents were evaluated for g_c, confirming the association between g_c and *ISTL1* in a haplotype-based association analysis. The findings of this study provide novel insights into the role of regulatory variants in the development of the maize leaf cuticle, and will ultimately assist breeders to develop drought-tolerant maize for target environments.

Introduction

The cuticle is a hydrophobic layer covering the epidermal surfaces of the shoot, which protects plants from dehydration, UV radiation, and pathogen attack (Shepherd and Wynne Griffiths, 2006; Xue et al., 2017). Cutin and waxes are the two major lipid components comprising the plant cuticle. Cutin is an insoluble matrix formed by extensive ester cross-linking of fatty acid derivatives and glycerol (Pollard et al., 2008; Fich et al., 2016). Soluble waxes, including alkanes, aldehydes, alcohols, ketones, and wax esters, are embedded within and on top of the cutin matrix (Yeats and Rose, 2013).

Cuticle composition and structure influence its water barrier function (Kerstiens, 2006). However, cuticle impermeability to water is not simply determined by wax load or cuticle thickness; instead, chemical composition and the organization of cuticle components appear to contribute to reducing non-stomatal water loss (Riederer and Schreiber, 2001). Thus, cuticle composition and structure are potentially relevant to plant drought tolerance. Mutants and transgenic plants with reduced wax load and compromised cuticle structure often show increased cuticular permeability and decreased drought tolerance (Zhou et al., 2013; Zhu and Xiong, 2013; Li et al., 2019), while overexpression of cuticle lipid biosynthesis enzymes or their transcriptional regulators can result in increased cuticular lipid abundance and drought tolerance (Aharoni et al., 2004; Zhang et al., 2005; Bourdenx et al., 2011; Wang et al., 2012). In wheat and barley, glaucousness, a visual manifestation of epicuticular wax crystals, was selected for during domestication (Bi et al., 2016; Hen-Avivi et al., 2016) and shows positive correlations with drought tolerance (Febrero et al., 1998; Guo et al., 2016; Busta et al., 2021). Collectively, these findings suggest a range of possibilities toward increased drought tolerance via cuticle modification in cereal crops.

Biological pathways involved in cuticle development affect its permeability to water in different ways. The formation of the plant cuticle requires not only the synthesis of cutin monomers and cuticular waxes but also the transport and polymerization of cutin monomers after transport (Yeats and Rose, 2013). Major pathways for the biosynthesis of cutin monomers and cuticular waxes have been elucidated in model plant systems (Yeats and Rose, 2013; Lee and Suh, 2015; Fich et al., 2016). A Golgi-mediated vesicle trafficking system delivers cuticle lipids from the intracellular membranes, where they are synthesized, to the plasma membrane (McFarlane et al., 2014) for export by ATP BINDING CASSETTE TRANSPORTER G (ABCG) family proteins (Pighin et al., 2004; Bird et al., 2007; Panikashvili et al., 2010; Bessire et al., 2011; Chen et al., 2011) and extracellular lipid transport proteins (DeBono et al., 2009; Kim et al., 2012). In addition, GLY-ASP-SER-LEU ESTERASE/LIPASE (GDSL lipase) is required for cutin polymerization in the developing cuticle (Girard et al., 2012; Yeats and Rose, 2013), and the degree of pectin esterification is critical to cuticle structure by facilitating the diffusion of cutin precursors (Bakan and Marion, 2017; Philippe et al., 2020). Transcription factors have also been identified to regulate various steps in cuticle development (Borisjuk et al., 2014; Elango et al., 2020), but how these transcription factors, biosynthesis and transport processes work together is not well understood.

Genome-wide association studies (GWAS) are widely used to identify genomic variants that are associated with a trait of interest (Tibbs Cortes et al., 2021). A genome-wide association study conducted by Lin et al. (2020) revealed that adult maize leaf cuticular conductance (g_c; defined as the rate of water loss when stomata are closed, an indirect measure of cuticular permeability to water) is likely controlled by large numbers of small effect alleles in maize, which agrees with the complex cuticle biosynthesis network and

transport mechanisms (Yeats and Rose 2013). Given that the GWAS of Lin et al. (2020) was limited by low mapping resolution and statistical power, the use of gene expression data with its offering of gene-level resolution and capturing of regulatory variation provides an opportunity to address such weaknesses. Therefore, transcriptome-wide association studies (TWAS), which identify significant associations between trait variation and transcript abundance across all expressed genes in a tissue, can provide complementary information for candidate gene identification in an independent test. Furthermore, integrating GWAS and TWAS results with an ensemble approach based on the Fisher's combined test (FCT) improves statistical power in candidate gene identification (Kremling et al., 2019), especially for leaf physiological traits with a complex genetic architecture (Ferguson et al., 2021; Pignon et al., 2021).

In this study, an integrated investigation was performed to dissect natural variation for leaf g_c in a maize inbred association panel scored with ~9 million SNP markers and transcript abundance from cuticle maturation zones (as identified by Bourgault et al. 2020) of adult leaves. Combining GWAS and TWAS results with a FCT boosted statistical power to identify genes for this complex trait. Evaluation of maize nested near-isogenic lines (nNILs) (Morales et al., 2020) confirmed the association between g_c and one of the identified plausible causal genes, which encodes a predicted INCREASED SALT TOLERANCE1-LIKE1 (ISTL1) protein.

Results

Relationship of g_c to cuticle composition and structure

Permeability of the leaf cuticle to water was measured as g_c on 310 genetically representative maize inbred lines from the Wisconsin Diversity (WiDiv) panel that were evaluated in two environments in San Diego, CA. To investigate how this phenotype of

moderately high heritability ($\hat{h}_l^2 = 0.73$) related to cuticle characteristics, we quantitatively analyzed selected features of adult leaf cuticle composition and structure in a subset of 51 WiDiv lines with highest and lowest g_c values. Our structural analysis focused on features of guard cell cuticles that could be analyzed quantitatively in a large number of samples via confocal microscopy, and seemed likely to affect the g_c trait. As illustrated in Supplemental Figure 1, we observed cuticular flaps projecting from the inner surfaces of guard cell pores, as described in many plant species, often called "outer cuticular ledges" (e.g. Hunt et al., 2017), which have been proposed to reduce water loss by enhancing the seal of closed stomata (Wilmer and Fricker, 1996; Zhao and Sack, 1999; Kosma and Jenks, 2007). We also observed cuticular bumps lining stomatal pores, which appear to permit a "zippering" of closed stomata (Supplemental Figure 1). However, variation in these traits was only weakly associated with g_c ($|r| < 0.15$ for all measures; Supplemental Table S1). Numerous studies have demonstrated relationships between cuticular wax composition and cuticle permeability (Zhou et al., 2013; Zhu and Xiong, 2013; Li et al., 2019), and wax composition was also found to vary in these extreme g_c lines. Therefore, we also investigated how variation in g_c may be related to wax composition. Significant associations were observed between g_c and a subset of fatty acids, wax esters, aldehydes, and alicyclics ($r = -0.299$ to -0.488; Supplemental Table S1).

To further query the relationship of g_c to cuticular wax composition and structural features, random forest analysis was conducted with five-fold cross-validation to determine the degree to which g_c can be predicted by the analyzed features, and which features are the strongest drivers of predictive ability. As shown in Figure 1 and Supplemental Table S1, we found that the abundance of several long chain wax esters and aldehydes (especially C54 wax esters and C30 aldehydes) to be important predictors of g_c. The overall predictive ability for

6

g_c from all measured characteristics combined was 0.35, with most of the predictive power coming from measurements of C54 wax esters and C30 aldehydes. The relationship of g_c to C54 wax ester abundance is interesting in relation to the findings of Bourgault et al. (2020), showing that acquisition of mature water barrier function during cuticle development in adult maize leaves coincided with a shift from alkanes to esters with carbon chain lengths > 49 as the most abundant wax class. Thus, our analysis of wax composition in extreme lines provides further evidence of a role for wax esters in protecting adult maize leaves against water loss, consistent with recent findings in Arabidopsis (Patwari et al., 2019). However, the majority of g_c is not explained by the cuticle features we analyzed, underscoring the complexity of this trait that merits further genetic examination through the combination of GWAS and TWAS.
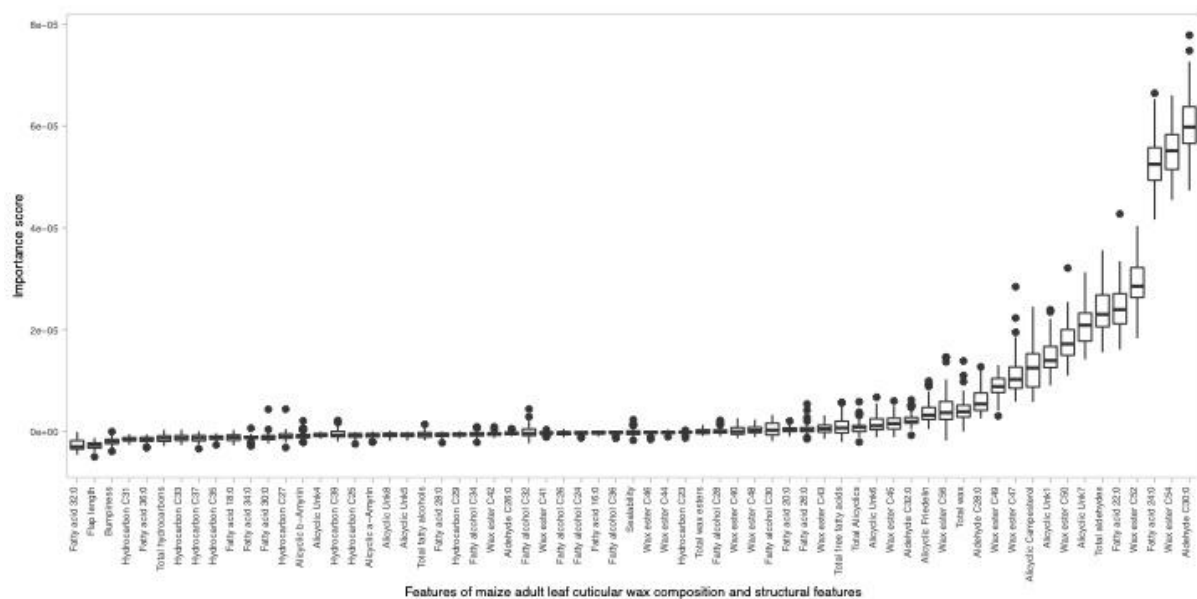

Figure 1. Box-plot of importance scores for cuticular wax composition and structural features in random forest regression to predict maize leaf cuticular conductance (g_c) using five-fold cross validation for 50 times.

Quantitative genetic analysis for identification of candidate genes for g_c

To improve statistical power in the identification of candidate genes for g_c beyond that achieved in GWAS alone by Lin et al. (2020), we performed the Fisher's combined test (FCT) to integrate GWAS and TWAS results (see Methods for more detail). A total of 22 plausible candidate genes were selected based on predicted functions that could be connected to cuticles from a total of 319 candidates (top 1%) from FCT (Supplemental Table S2), 273 (top 0.01%) from GWAS (Supplemental Tables S3 and S4), and 200 (top 1%) from TWAS (Supplemental Table S5). Of these 22 genes, 14 were identified by FCT, five by TWAS and three by GWAS; nine were identified by two or more of these strategies (Table 1; Supplemental Figure S2). Supplemental Table S6 lists the closest Arabidopsis and rice relatives of these 22 genes.

Table 1. Plausible candidate genes identified in a genome-wide association study (GWAS), a transcriptome-wide association study (TWAS) and the Fisher's combined test (FCT) for leaf cuticular conductance (g_c) in the Wisconsin diversity panel.

Gene ID	Annotation	Chromosome	Gene start	Gene end	Functional category[b]	Method
Zm00001d029141	MIEL1-like	1	60,322,950	60,326,019	RCD	GWAS, FCT
Zm00001d032788	GDSL esterase/lipase	1	238,190,770	238,192,464	CBE	TWAS, FCT
Zm00001d002843	MYB54	2	24,293,846	24,295,521	RCD	FCT
Zm00001d005087	GLK35	2	157,736,007	157,737,776	RCD	TWAS, FCT
Zm00001d039411	α1-COP	3	3,708,911	3,716,373	IMT	TWAS, FCT
Zm00001d040788	SEC15B	3	65,988,895	65,991,327	IMT	GWAS
Zm00001d042612	CER7	3	173,571,381	173,573,177	RCD	FCT
Zm00001d043509	PME40	3	202,333,699	202,335,784	CWB	TWAS
Zm00001d044162	WRKY64	3	220,824,499	220,827,409	RCD	TWAS
Zm00001d048953	KCS3	4	10,603,315	10,607,553	CBE	TWAS, FCT
Zm00001d049479	ISTL1	4	31,729,260	31,732,664	IMT	GWAS, TWAS, FCT
Zm00001d050185	MYB108	4	71,478,007	71,479,474	RCD	TWAS
Zm00001d051599	Rab-GAP/TBC	4	164,071,426	164,090,678	IMT	TWAS, FCT
Zm00001d012964	SEC14	5	2,431,015	2,437,696	IMT	FCT
Zm00001d015477	CER3-like	5	92,669,605	92,675,745	CBE	TWAS, FCT
Zm00001d036765	CER9	6	100,099,887	100,107,565	RCD	TWAS
Zm00001d038404/ Zm00001d038405[a]	Ypt/Rab-GAP	6	156,394,901/ 156,409,880	156,403,240/ 156,412,505	IMT	GWAS
Zm00001d022364	SNARE	7	176,149,978	176,160,500	IMT	TWAS, FCT
Zm00001d010426	CER5/ABCG12-like	8	114,527,255	114,533,597	CBE	TWAS
Zm00001d012175	PAE5	8	169,547,499	169,551,527	CWB	FCT
Zm00001d025180	Ypt/Rab-GAP	10	107,999,211	108,003,428	IMT	FCT

[a]In B73 RefGen_v4, this gene was erroneously split into the two gene models as indicated, but in B73 RefGen_v5 they have been correctly merged into a single gene model with identifier Zm00001eb289920.
[b]RCD, regulator of cuticle development; CBE, cuticle biosynthesis and export; IMT, intracellular membrane trafficking; CWB, cell wall biosynthesis.

Of the 14 genes identified by FCT, one (*Zm00001d049479*) encoding the closest

maize relative of INCREASED SALT TOLERANCE1-LIKE1 (ISTL1) in Arabidopsis, which

functions in multivesicular body formation within endosomes (Buono et al,. 2016) was also

identified by both TWAS and GWAS, and was the only candidate gene identified by all three

methods (Table 1; Figure 2). Additionally, five of the 14 genes identified by FCT encode

proteins with predicted functions in intracellular membrane trafficking, three of which were

also identified by TWAS (Table 1). Two of these five, *Zm00001d039411* and

Zm00001d012964, encode proteins with sequence identity to α1 COAT PROTEIN (α1-COP)

and SEC14 in Arabidopsis, respectively, which function in vesicle formation at the Golgi

surface (Jouannic et al., 1998; Lee and Goldberg, 2010). *Zm00001d051599* and

Zm00001d025180 encode proteins with Rab GTPase-activating protein (Rab-GAP) domains,

predicted to function in regulation of vesicle transport (Stenmark, 2009), and

Zm00001d022364 encodes a soluble NSF attachment protein receptor (SNARE) homolog

with an expected function in vesicle fusion with target membranes (Chen and Scheller, 2001).
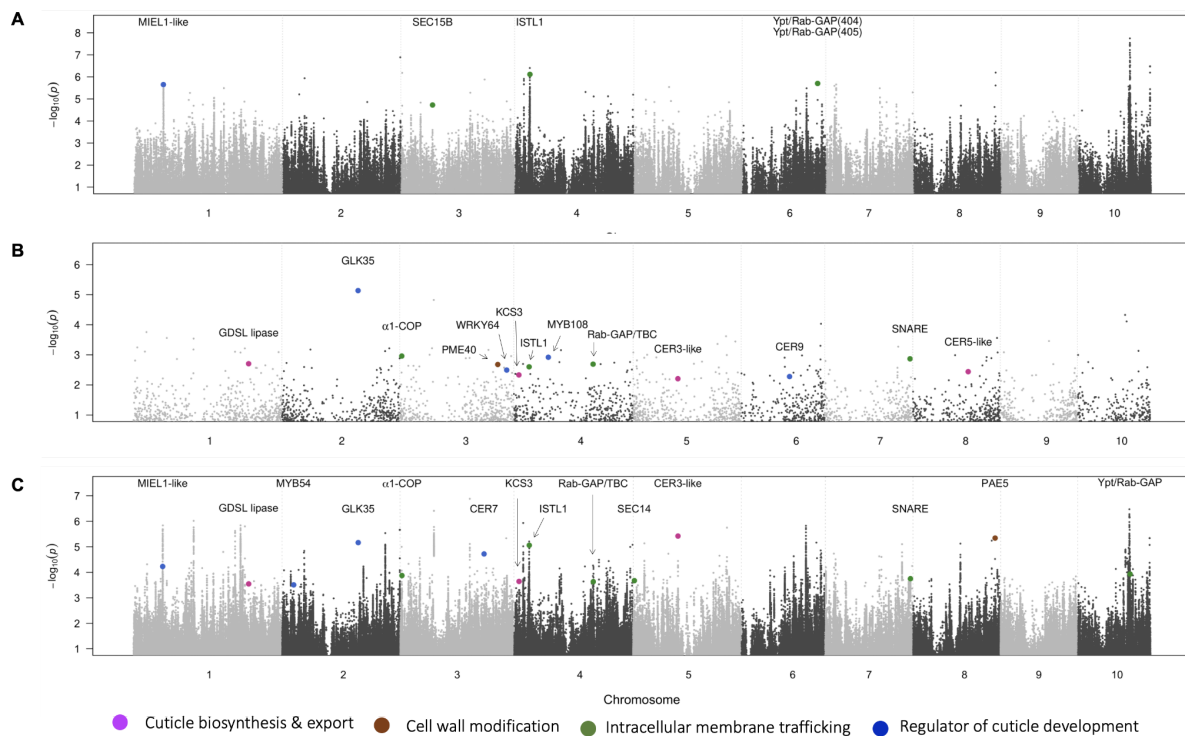
Figure 2. Manhattan plot of results from a genome-wide association study (A; GWAS), a transcriptome-wide association study (B; TWAS), and Fisher's combined test (C; FCT) of adult maize leaf cuticular conductance (g_c) across two environments in San Diego, CA. (A) The -$\log_{10}$ P-value of each SNP tested in a mixed linear model analysis of g_c is plotted as a point against its physical position (B73 RefGen_v4) for the 10 chromosomes of maize. The peak SNPs closest to plausible candidate genes are colored based on their functional categories. (B) The -$\log_{10}$ P-value of each gene transcript tested in a mixed linear model analysis of g_c is plotted as a point against its physical position (B73 RefGen_v4). Plausible candidate genes are colored based on their functional categories. (C) The -$\log_{10}$ P-value of each of the top 10% SNPs in GWAS paired with the nearest gene in FCT plotted as a point against its physical position (B73 RefGen_v4). The most significant SNP-gene pairs related to plausible candidate genes identified in the FCT are colored based on their functional categories.

The remaining eight plausible candidate genes identified by FCT, four of which were also identified by TWAS and one also by GWAS, encode proteins with a variety of functions. Two (*Zm00001d048953* and *Zm00001d015477, 3-KETOACYL-COA SYNTHASE 3* [*KCS3*] and *GLOSSY1-like/ECERIFERUM3-like* [*GL1-like/CER3-like*], respectively), are homologs of Arabidopsis genes with functions in cuticular wax biosynthesis (Lee et al., 2009; Bernard et al., 2012). A third (*Zm00001d032788*) encodes a predicted GDSL esterase/lipase gene closely related to Arabidopsis and tomato cutin synthases (Girard et al., 2012; Yeats et al.,

10

2012). Two FCT-identified genes (*Zm00001d005087* and *Zm00001d002843,* annotated as GLK35 and MYB54, respectively), encode transcription factors with potential functions in regulation of cuticle biosynthesis (Kant et al., 2008; Dubos et al., 2010; Oshima and Mitsuda, 2013). *Zm00001d029141* is a homolog of Arabidopsis MIEL1 (MYB30-INTERACTING E3 LIGASE 1)*,* which regulates the stability of two MYB family regulators of wax biosynthesis (Marino et al., 2013; Gil et al., 2017). *Zm00001d042612* encodes a homolog of Arabidopsis CER7, a core subunit of the RNA degrading exosome associated with alkane biosynthesis (Hooker et al., 2007). The final FCT-identified gene listed in Table 1, *Zm00001d012175* (PAE5), encodes a predicted pectin-modifying enzyme.

A total of five additional plausible candidate genes were uniquely detected by TWAS (Figure 2; Table 1). *Zm00001d043509* encodes a putative pectin methylesterase (PME40). *Zm00001d010426* encodes a homolog of Arabidopsis CER5/ABCG12, which functions in cuticular wax secretion (Pighin et al., 2004). *Zm00001d036765* encodes a homolog of Arabidopsis CER9, which regulates the abundance of both cutin and cuticular waxes (Lü et al., 2012). The two remaining TWAS-identified genes (*Zm00001d050185* and *Zm00001d044162*) encode homologs of Arabidopsis WRKY and MYB family transcription factors with potential roles in cuticle development, which are differentially expressed across maize leaf cuticular developmental stages (Qiao et al., 2020).

GWAS uniquely identified three additional plausible candidate genes predicted to be involved in the intracellular membrane trafficking system (Figure 2; Table 1). Two adjacent genes (*Zm00001d038404* and *Zm00001d038405*) putatively encode Ypt/Rab-GAPs, regulators of vesicle fusion (Stenmark, 2009). *Zm00001d040788* encodes a homolog of Arabidopsis SEC15B, an exocyst subunit that activates vesicle fusion with the plasma membrane during polarized secretion (Guo et al., 1999; Mayers et al., 2017).

11

Haplotype-based association analysis for g_c in the WiDiv panel

The *ISTL1* gene emerged from our analysis as a strong candidate regulator of g_c, since it was identified in our prior GWAS (Lin et al., 2020) and the only gene detected by all three analyses (GWAS, TWAS and FCT) in this study. To further investigate the association between g_c and genetic variation within the WiDiv population in the vicinity of *ISTL1*, we analyzed the association between g_c and haplotypes at 91 haploblocks (blocks of haplotypes) within 200 kb of *ISTL1*. Of these, Block 65 was the most significantly associated with g_c (*P*-value 4.07×10^{-7}) in the WiDiv panel (Figure 3; Supplemental Table S7), having three haplotypes defined by three SNPs including the peak GWAS SNP (4-30231047) only 1,219 bp away from the 3′ end of *ISTL1*. The second most strongly associated (*P*-value 4.07×10^{-7}) haploblock, Block 52 (Figure 3; Supplemental Table S7), consisted of 10 haplotypes, with the 16 defining SNPs spanning a 8.6 kb region more than 20 kb upstream of *ISTL1*. These findings suggest that the underlying causal variants are located both upstream and downstream of the *ISTL1* gene.
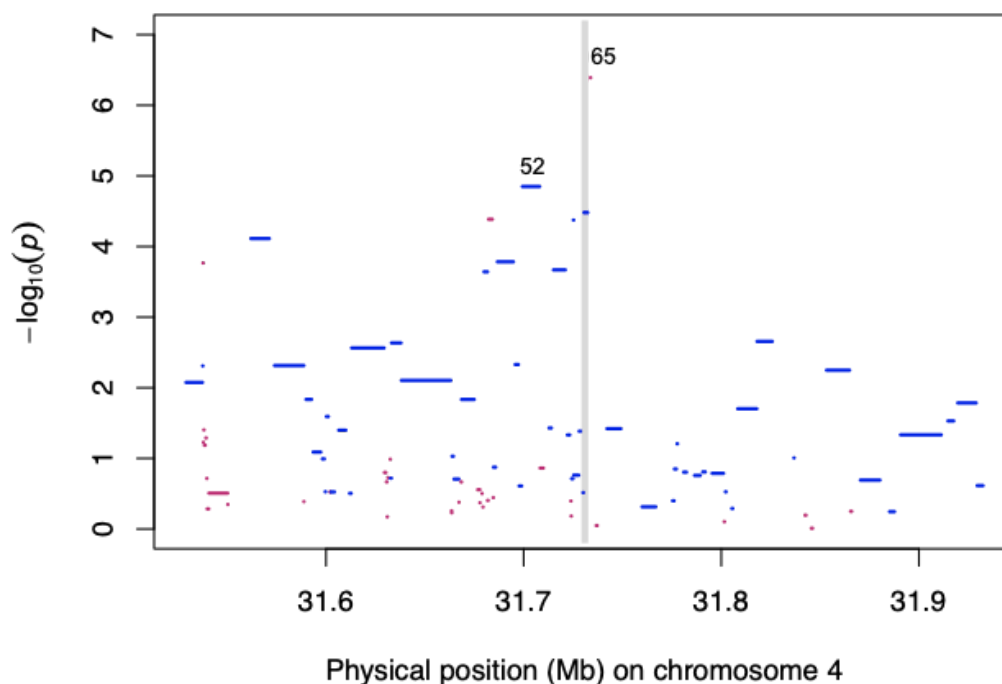
12

Figure 3. Manhattan plot of results from a haplotype-based association analysis in the vicinity of *ISTL1* for adult maize leaf cuticular conductance (g_c) combined from two environments in San Diego, CA, in the Wisconsin diversity panel. The -$\log_{10}$ P-value of each haploblock tested in a mixed linear model analysis of g_c is plotted as a short horizontal line against its physical position (B73 RefGen_v4). The length of a short horizontal line corresponds to the size of that haplotype block. The grey vertical bar represents the physical position of *ISTL1*. The top two haploblocks (Blocks 65 and 52) that are most associated with g_c are indicated. Polymorphic and monomorphic haploblocks among the evaluated nested near-isogenic lines are colored in blue and red, respectively.

Validation of *ISTL1* haplotype effects with nested NILs

A set of 12 nested near-isogenic lines (nNILs) (Morales et al., 2020) containing the *ISTL1* genomic region from diverse inbreds that are mostly not represented in the WiDiv panel, introgressed into B73 (Supplemental Table S8), were used to validate haplotype effects observed at *ISTL1* in the WiDiv panel. All 12 evaluated nNILs had lower g_c relative to their corresponding donor parent (Figure 4; Supplemental Table S9), which is to be expected given that the recurrent parent, B73, had a comparable or lower g_c than the donor parents. Compared to B73, two NILs with Tzi8 introgressions (Tzi8-B73_NIL_1160

13

and Tzi8-B73_NIL_1308) and one NIL with an Oh43 introgression

(Oh43-B73_NIL_1005) had significantly lower g_c, while four of the donor parents (Ki11,

NC350, Oh43 and Tx303) showed significantly higher g_c (Figure 4; Supplemental Table
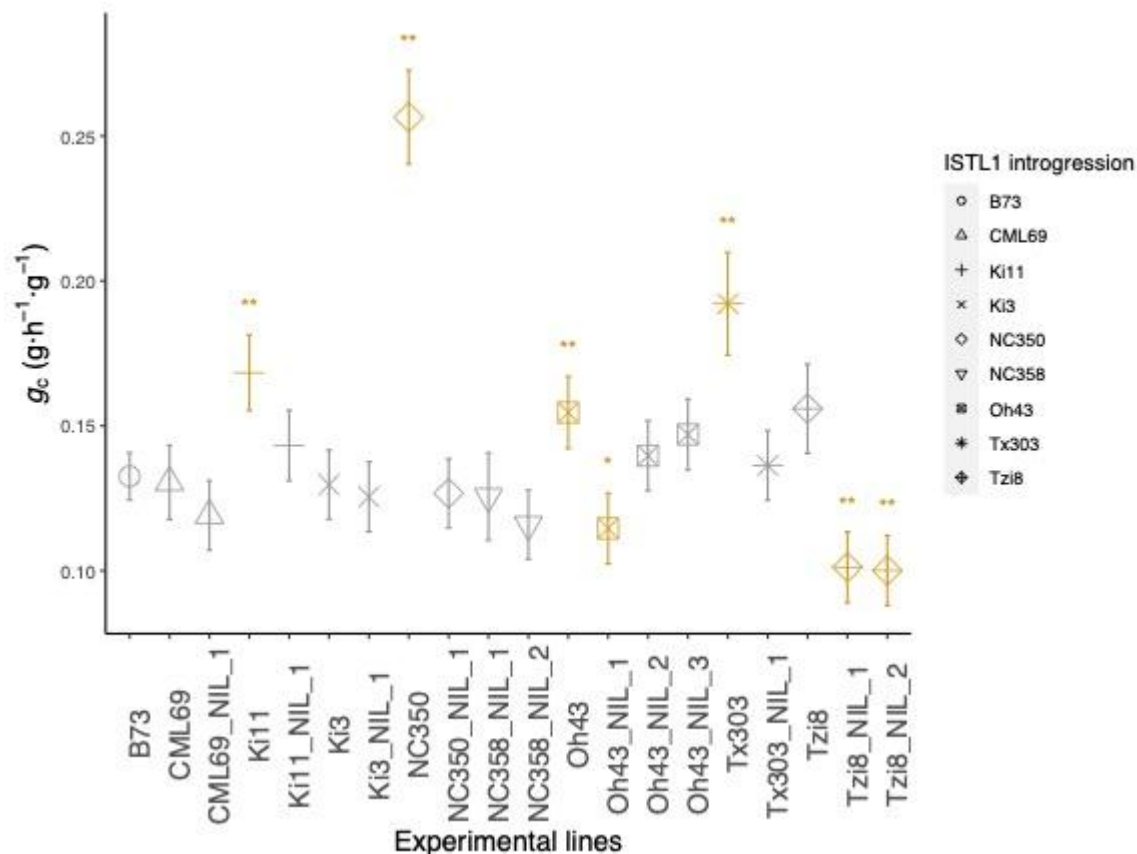
S9).

Figure 4. Scatter plot of best linear unbiased estimator values for adult maize leaf cuticular conductance (g_c) for B73, nested near-isogenic lines (nNILs) containing *ISTL1* introgressions and their donor parent lines across two environments in San Diego, CA, in 2020. Lines with g_c significantly different from B73 are highlighted in orange. '*' represents P-value < 0.1 and '**' represents P-value < 0.05 after adjustment by the Dunnett's method.

Of the two most strongly associated haploblocks in the WiDiv panel, only Block

52 (Supplemental Figure S3A), the second ranked haplotype-based association, was

polymorphic among the nNILs (Figure 3). This haploblock, consisting of three segregating

haplotypes (Supplemental Figure S3B), had a significant association (P-value < 0.020)

14

with g_c among nNILs. The haplotype effect of the Tzi8 NILs, which possess a unique haplotype relative to B73 and the other donor parents, was significantly (P-value < 0.05) different from the other two haplotypes (Supplemental Figures S3B). The Tzi8 haplotype had an effect of 0.1007 g·h^{-1}·g^{-1}, whereas the other two haplotypes averaged ~0.13·h^{-1}·g^{-1}. The smaller effect of the Tzi8 haplotype could explain why both Tzi8 nNILs had the lowest observed g_c values (Figure 4). Haplotype effects of Block 52 strongly correlated (r = 0.924) with those estimated in the WiDiv panel (Supplemental Figure S3C), indicating that the haplotype effects at *ISTL1* for g_c could be validated in an independent panel having a different population structure.

Discussion

The plant cuticle, as a crucial barrier to water loss (Shepherd and Wynne Griffiths, 2006; Xue et al., 2017), is potentially related to plant drought tolerance (Aharoni et al., 2004; Wang et al., 2012; Zhou et al., 2013; Li et al., 2019). Although major pathways for wax and cutin monomers biosynthesis have been elucidated in model species (Yeats and Rose, 2013; Bakan and Marion, 2017), genetic controls of water loss through leaf cuticles, i.e. g_c, remain largely unknown. In a prior study (Lin et al., 2020), we used 235,004 GBS SNP markers for a GWAS utilizing 451 lines of the WiDiv panel to identify five genomic regions impacting g_c. In the present study, we extended this earlier work by imputing genotypes for ~9 million SNPs and evaluating the following two features for a subset of 310 WiDiv lines in two growth environments: g_c, and transcript abundance for 20,013 genes in the cuticle maturation zones of developing adult leaf blades. These data were analyzed by GWAS, TWAS, and FCT, greatly increasing the power to identify genes with likely roles in determination of g_c. A total of 22 high-scoring genes (discussed below) from one or more of these tests were identified as plausible candidates based on predicted

15

functions potentially impacting g_c via the regulation of cell wall biosynthesis (2 genes), cuticle biosynthesis and export (4 genes), intracellular membrane trafficking (9 genes) and cuticle development (7 genes) (Figure 2; Table 1). Of the 23 GWAS peaks detected in this study, only the GWAS signal associated with *ISTL1* was also found in our previous study that identified five genomic loci associated with g_c (Lin et al., 2020). The main reason for such limited overlap is most likely the extreme difference in the growth environment where associations were found in the prior study (Maricopa, AZ which is very hot and dry) compared to that used for the present study (San Diego, CA, which is relatively cool and humid). Subtle changes in population structure when subsetting the WiDiv population, and different criteria for calling GWAS peaks in this study, may also have contributed to differences in GWAS signals.

Two of the genes identified as plausible candidates for g_c determination have predicted functions in cell wall modification, in particular pectin esterification (PAE5 and PME40). Cutin-embedded polysaccharides present a high degree of esterification, including methylation and acetylation (Philippe et al., 2020). Pectin esterification could be critical to cuticle structure by favoring interactions between cutin and cutin-embedded polysaccharides as well as facilitating the diffusion of cutin precursors (Bakan and Marion, 2017; Philippe et al., 2020) potentially modulating cuticle function as a barrier to water loss.

Our study also pointed to genes encoding predicted cuticle lipid biosynthetic enzymes. FCT identified a maize homolog of *CUS1/LTL1*, encoding a GDSL acyltransferase that polymerizes cutin precursors extracellularly in Arabidopsis and tomato (Girard et al., 2012; Yeats et al., 2012). FCT and TWAS both identified maize KCS3, a putative component of the fatty acid elongase (FAE) complex that extends acyl chains

16

produced by the plastid to very-long-chain fatty acids. Its closest Arabidopsis relative, KCS2, functions in the two-carbon elongation of acyl chains to C22 very long chain fatty acyl precursors required for cuticular wax and root suberin biosynthesis (Franke et al., 2009; Lee et al., 2009). Notably, several *KCS* homologs are upregulated during cuticle maturation in the developing maize adult leaf epidermis (Qiao et al., 2020). TWAS and FCT also identified a maize homolog of *AtCER3* and *ZmGL1*, genes involved in alkane biosynthesis (Sturaro et al., 2005; Bernard et al., 2012). However, this *GL1/CER3*-like maize gene has not been previously characterized by forward or reverse genetic approaches. Finally, TWAS analysis identified a maize homolog of Arabidopsis *ABCG12/CER5*, encoding a half-transporter of the ABCG family that forms heterodimers with ABCG11 to transport cuticle lipids across the plasma membrane (Pighin et al., 2004; McFarlane et al., 2010). Unlike the other cuticle biosynthesis candidates highlighted in this study, this maize *CER5* homolog is epidermally upregulated during cuticle maturation (Qiao et al., 2019). Identification of several candidate genes with predicted functions in the biosynthesis and transport of cuticle lipids is consistent with our finding that g_c can be partially predicted by cuticular wax composition.

The biosynthesis of plant cuticles is regulated in a complex network (Yeats and Rose, 2013). In this study, we identified maize homologs of three known regulators of cuticle development (Figure 2; Table 1). This includes a maize homolog of Arabidopsis CER7, a core subunit of exosomal 3′-to-5′ exoribonuclease. This enzyme reduces the degradation of CER3/WAX2/YRE transcripts (Hooker et al., 2007). CER3 and CER1 proteins interact to jointly synthesize VLC-alkane (Bernard et al., 2012), an important class of cuticle waxes that is related to drought resistance (Kosma et al., 2009; Wang et al., 2015; Wang et al., 2020). Interestingly, a different maize CER7 homolog was associated

17

with g_c in our prior GWAS (Lin et al., 2020), underscoring the potential importance of CER7 proteins in the regulation of g_c. TWAS identified maize CER9, homologous to an Arabidopsis E3 ubiquitin ligase whose deficiency enhances plant drought tolerance via alteration of cuticle composition and reduction of cuticular transpiration (Lü et al., 2012). Finally, both GWAS and FCT identified a maize homolog of Arabidopsis MIEL1, an E3 ligase that regulates ABA sensitivity by controlling the protein stability of MBY96 and MYB30, R2R3-MYB subfamily regulators of extracellular lipid biosynthesis (Raffaele et al., 2008; Seo et al., 2011), which are important for balanced cuticular wax biosynthesis (Marino et al., 2013; Lee and Seo, 2016; Gil et al., 2017).

Four putative transcriptional regulators of cuticle biosynthesis were also identified as plausible g_c candidates: one WRKY, one homeodomain, and two MYB transcription factors (Figure 2; Table 1). *WRKY64* has not been directly associated with cuticle development, but a closely related gene in Arabidopsis (*AtWRKY28; mybAT4G18170*) (Supplemental Table S6) (Jiang et al., 2014) and its maize homolog, *ZmWRKY28* (*Zm00001d011413*), are implicated in cuticle development (Qiao et al., 2020). MYB108 is homologous to ATMYB36, which directly and positively regulates formation of the Casparian strip (characterized by deposition of suberin, which is chemically similar to cutin) (Kamiya et al., 2015). MYB108 is also differentially expressed across the time course of maize leaf cuticular maturation (Qiao et al., 2020). The closest Arabidopsis relative (ATMYB17) of FCT-identified MYB54 (Supplemental Table S6) belongs to the same clade as two well-characterized MYBs of cuticle biosynthesis, MYB106 and MYB16 (Dubos et al., 2010; Oshima and Mitsuda, 2013). GLK35, identified via TWAS and FCT, is a predicted homeodomain-containing transcription factor that has not been previously studied, but its Arabidopsis homolog (AT2G40260) showed > 8-fold change in

transcript abundance in response to heat stress (Kant et al., 2008). Therefore, GLK35 could be a regulator for g_c induced by high temperatures. Together, these candidate genes highlight the potential for developmental and environmental regulation of cuticle formation through transcriptional, post-transcriptional, and post-translational mechanisms.

Several candidate genes identified in our study have predicted functions in intracellular membrane trafficking, the process by which proteins and lipids are moved from one cellular location to another. These candidates are noteworthy because of prior evidence of a role for membrane trafficking in cuticle formation (McFarlane et al., 2014). Two of our candidate genes encode proteins predicted to function in vesicle formation. One of these is a putative COPI coat component (α1-COP). COPI has not been widely studied in plants, but COPI deficiency in yeast leads to protein accumulation in the ER and global secretion deficiency (Lodish et al., 2016). The other is a putative member of SEC14 family phosphatidylinositol transfer protein, and is different from the two SEC14 homologs we identified previously as candidate regulators of g_c via GWAS (Lin et al., 2020). Four of our candidate genes have predicted functions in vesicle targeting or fusion, including one putative SNARE, a putative exocyst subunit (SEC15B), and three Rab-GAP domain family proteins with expected functions in regulation of Rab GTPases. By different mechanisms, exocyst, Rab and SNARE proteins mediate docking and fusion of secretory vesicles, ensuring the specificity of fusion between particular vesicle types and their appropriate target membranes (Saito and Ueda, 2009). These candidate regulators of g_c have potential functions in targeted delivery, to the outer face of epidermal cells, of cuticular lipids as well as proteins such as lipid transporters that are needed for cuticle assembly, thereby impacting the water barrier function of the cuticle.

The final candidate with a predicted membrane trafficking function, ISTL1, is of particular interest because it was also identified in our prior GWAS (Lin et al., 2020), and is the only candidate in the present study identified by all three approaches: GWAS, TWAS, and FCT. Yeast IST1 promotes the formation of vesicles inside multivesicular bodies (MVBs; (Hill and Babst, 2012)), a function shared by Arabidopsis ISTL1 (Buono et al., 2016). The MVB is a late endosomal compartment where plasma membrane-derived proteins are degraded after internalization via endocytosis (Lodish et al., 2016). Recent work has demonstrated that Arabidopsis ISTL1, in combination with a functionally related protein LIP5, is needed to maintain an ABCG family transporter at the plasma membrane of anther tapetal cells, and promotes the accumulation of certain wax components deposited on pollen grains by tapetal cells (Goodman et al., 2021). Thus, it can be speculated that maize ISTL1 facilitates the localization of plasma membrane transporters that export cuticle components in maize epidermal cells. Alternatively, cuticular lipids may be delivered extracellularly by fusion of MVBs with the plasma membrane.

The association between g_c and the *ISTL1* genomic region was further verified in a haplotype-based association analysis, which has increased statistical power for detecting marker-trait associations relative to GWAS, and can also identify favorable haplotypes for breeding projects. Haplotype-based association tests were performed in both the diversity panel and among nNILs. The strong correlation for haplotype effects for Block 52 in the diversity panel and nNILs confirmed the genetic associations across genetic materials and environments (2018 vs 2020). However, the power of our nNIL analysis to detect genetic effects of haplotype variation in the *ISTL1* region was reduced by two limitations. First, we analyzed only 12 nNILs, which represented all the genetic variation in this collection (Morales et al., 2020) with non-B73 haplotype in the *ISTL1* region. This is a relatively

small population for this type of analysis when investigating a locus with a moderate genetic effect for a polygenic trait. In addition, Block 65, which showed the strongest association with g_c in the diversity panel, was monomorphic among nNILs, so the phenotypic effects of this haploblock could not be investigated in the nNIL analysis. While these experiments provided modest additional evidence of the impact of haplotype variation at *ISTL1* on g_c, a single gene mutation analysis would be ultimately needed to prove the role of this gene.

In summary, GWAS, TWAS and FCT together revealed 22 genes and regulatory factors associated with maize adult leaf g_c. These genes have predicted functions related to the synthesis and function of the cuticle. A candidate causal gene encoding ISTL1 was associated with g_c by all three tests. Haplotype-based association analysis using maize nNILs containing introgressions of the *ISTL1* genomic region from eight inbred donors into a common background (B73) corroborated marker-trait associations found in WiDiv lines, providing additional evidence of a role for *ISTL1* as a determinant of g_c. The *ISTL1* region haplotypes harbored by Tzi8 were consistently associated with reduced g_c, and could be introgressed into adapted maize germplasm to potentially increase drought tolerance via a reduction in non-stomatal transpiration (evaporation across the cuticle).

21

Materials and Methods

Plant materials and experimental design

A set of 323 maize inbred lines from the Wisconsin Diversity (WiDiv) panel (Hansey et al., 2011) was planted as a single replicate in a 17×19 incomplete block design at two different times (May and June 2018) in the same field (WiDiv18) at University of California San Diego, San Diego, CA, for evaluation of adult leaf cuticular conductance (g_c). The set consisted of 89 lines that had been previously shown to have the highest or lowest g_c (extreme g_c lines) in San Diego field experiments (Lin et al., 2020) in addition to 234 lines with restricted flowering time selected for maximal genetic diversity using the CDmean method (Rincent et al., 2012). The field design in both environments was augmented by including Mo17 within each incomplete block. Experimental units were one-row plots of 3.66 m in length having approximately 12 plants with 1.02 m inter-row spacing and a 0.61 m alley.

g_c evaluation and analysis

To inform sampling times for each plot in the WiDiv18 experiment, flowering time (days to anthesis, DTA) was scored as previously described (Lin et al., 2020). The method of Lin et al. (2020) was used to evaluate g_c on an adult leaf from the primary ear node (or one leaf immediately above/below) of five plants (no fewer than two if five were unavailable) from each plot at flowering. The calculation of adult leaf cuticular conductance (g_c) from unit surface area was as follows:

$$g_c \ (\text{g·h}^{-1}\text{·g}^{-1}) = -\ b\ /\ \text{dry weight,}$$

where b (g·h^{-1}) is the coefficient of the linear regression of leaf wet weight (g) on time (h), and dry weight (g).

For each phenotype (g_c or DTA), a mixed linear model was fitted in ASReml-R version 3.0 (Gilmour et al., 2009) that had grand mean and check as fixed effects and

22

environment, genotype, genotype-by-environment (G×E) interaction, incomplete block within environment, and column within environment as random effects. Model generated Studentized deleted residuals (Neter et al., 1996) were used to remove significant outliers (Bonferroni $\alpha = 0.05$). The model was refitted for each outlier-screened phenotype to generate variance components to calculate heritability on a line-mean basis according to Lin et al. (2020). An iterative mixed model fitting procedure in ASReml-R version 3.0 was then used to select a best-fit model for each outlier-screened dataset (Supplemental Table S10) to generate a best linear unbiased predictor (BLUP) for g_c and DTA for each line (Supplemental Table S11) following the approach of Lin et al. (2020).

Phenotypic and statistical analysis of extreme g_c lines

Of the extreme g_c lines evaluated in the WiDiv18 experiment, we eliminated six classified as flint, popcorn, or tropical, resulting in 51 lines mostly having membership in the NSS or SSS subpopulations (Lin et al., 2020) that were examined via super-resolution confocal microscopy. For each line, 2-5 plants from each environment were analyzed by visualizing the cuticle with lipid stain Fluorol Yellow, and images for multiple stomata were collected for each plant ($n = 2\text{-}16$ stomata per plant). Tissue samples excised from the middle portion of adult leaf blades were collected and processed with Fluorol Yellow staining to visualize cuticle features as described by Matschi et al. (2020). Collected images were processed through superresolution Airyscan, and composite pictures were processed through ImageJ. For stomatal cuticle analysis, stomata were chosen where the section presented the middle third of a stomate (where the cross sectional area of subsidiary cells was at least equal to that of the guard cells). Using ImageJ, four cuticle phenotypic features were calculated as illustrated in Supplemental Figure S1: bumpiness, cuticular flap length, roundness, and sealability. For each of the four cuticle features, means of stomatal measurements were

23

calculated for each individual plant. Roundness was not statistically analyzed because of its very low genetic variation.

For the same 51 lines, waxes were extracted by submerging the mature leaf tissue in pure chloroform for 60 s. Extracted leaf wax samples were evaporated under a gentle stream of nitrogen, and wax samples were analyzed with gas chromatography as described in Bourgault et al. (2020).

Random forest regression models were fit to assess the prediction accuracy of g_c using 62 wax and three cuticle features. The 65 features used in modelling were calculated as an across-environment average of values that had been screened for outliers by Studentized deleted residuals (Supplemental Table S12). Forests were grown with the 'cforest' function from the R package 'party' version 1.3-7 (Strobl et al., 2007). Fivefold cross-validation was performed 50 times to evaluate the mean predictive accuracy for g_c as described in (Lin et al., 2020). Variable importance measures were calculated using the 'varimp' function in the 'party' package. Parameters used for the number of trees grown (*ntree*=1000) and predictors sampled (*mtry*=10) were those that maximized predictive accuracy.

Construction of genomic datasets for association analyses

Genotypic data

The genotype data processing and imputation approach of Wu *et al*. (2021) was implemented to generate a SNP marker set in B73 RefGen_v4 coordinates for the WiDiv panel. Briefly, a target SNP set consisting of 159,878 biallelic SNPs scored on 451 WiDiv lines with all heterozygous genotypes set to missing was generated by filtering raw genotypes (exclude singletons and doubletons, call rate $\geq$ 20%, heterozygosity $\leq$ 10%, and index of panmixia $\geq$ 0.8) of a 955,690 genotyping-by-sequencing (GBS) SNP set obtained from Lin et al. (2020). The reference SNP genotype set of Wu *et al*. (2021), which consisted of

14,613,169 SNPs derived from maize HapMap 3.2.1 (Bukowski et al., 2018), was imputed based on the target GBS SNP set in the 451WiDiv lines via BEAGLE v5.0 (Browning et al., 2018). The resultant set of imputed genotypes was filtered to enhance SNP quality for the 310 lines with g_c BLUP values, producing a set of 9,715,072 biallelic SNPs with MAF $\geq$ 5% and predicted dosage r^2 (DR2) $\geq$ 0.80 for conducting GWAS (Supplemental Data Set 1).

RNA sequencing data

Leaf samples for RNA sequencing (RNA-seq) were collected from plants of the WiDiv18 experiment that had not been destructively sampled for evaluating g_c. In both environments, three plants at a similar stage of development per plot were selected to collect tissue samples between 9-11 am PST from the proximal, immature and actively growing 10-30% section (developing cuticle) of the total length of the longest unexpanded adult leaf (sheath length < 2 cm). The sampled leaf sections, which were immediately frozen in liquid nitrogen, were from the second, third, or fourth fully adult leaf based on the last leaf with epicuticular wax reported by Hirsch et al. (2014). For each plot, leaf tissue samples from the three plants were ground on dry ice and equally pooled to form a composite ~100 mg tissue sample. All samples were randomized into 96-well plates for RNA extraction with the Direct-zol-96 RNA Kit (Zymo Research, Irvine, CA). Libraries constructed with the Lexogen QuantSeq 3′ mRNA-Seq Library Kit FWD (Lexogen, Greenland, NH) were sequenced on an Illumina NextSeq 500 (Illumina, San Diego, CA) at the Genomics Facility of the Cornell Institute of Biotechnology.

Raw 3′ QuantSeq reads were cleaned by trimming Illumina adaptors, the first 12 bases, and polyA tails in accordance with Lexogen recommendations. Cleaned reads were aligned to the B73 RefGen_v4 reference genome (Jiao et al., 2017) using HiSAT2 version 2.1.0 (Kim et al., 2019) with --rna-strandness F and other default parameters. Next, counts

25

were generated in HTSeq version 0.11.2 (Anders et al., 2015) using B73 version AGPv4.42

annotation with --format=bam, --type=gene, and --stranded=yes, followed by normalization

of the count data using the DESeq2 rlog function (Love et al., 2014). All genes with a

normalized count of less than or equal to zero in all samples were removed.

To verify sample provenance, SNPs were called using the 3′ QuantSeq read

alignments and compared to GBS SNPs (Lin et al., 2020). SNPs calls were generated from

a pileup file created using SAMTools mpileup with default parameters. The called 3′

QuantSeq SNPs were filtered for quality as previously described in (Lin et al., 2020), and

they were used to calculate percent identity between each pair of 3′ QuantSeq and GBS

samples. With a neighbor-joining tree constructed from a percent identity-based

dissimilarity matrix in R version 3.5.1 (R core team, 2018), RNAseq samples (14 in Env1

and 7 in Env2) not clustered with their corresponding GBS samples were excluded. Dataset

quality was further improved by removing RNAseq samples (18 in Env1 and 28 in Env2)

with < 2 million or > 20 million reads, or an overall alignment rate < 60%. The final dataset

consisted of 292 (280 unique lines) samples for Env1 and 305 (288 unique lines) samples

for Env2.

A mixed linear model was used to combine transcripts that had non-zero

rlog-transformed counts for at least 25% of lines in each environment. To generate BLUPs

for the retained 20,018 transcripts, the fitted model was the same as that used for generating

phenotypic BLUPs, with the exception that sequencing lane was also included as a random

effect. To account for inferred confounders, the probabilistic estimation of expression

residuals (PEER) (Stegle et al., 2010) approach was applied to the 310 line × 20,018 gene

matrix of BLUP expression values. The resultant PEER values after extracting 20 learned

factors were screened for outliers by Studentized deleted residuals. Transcripts with

26

missing values in over 10% of the population after outlier removal were filtered out,

producing a final set of 20,013 transcripts across 310 lines for conducting TWAS

(Supplemental Data Set 2).

Association analyses

GWAS

Each of the 9,715,072 SNPs was tested for an association with BLUP values of g_c

from the 310 lines with a mixed linear model (Zhang et al., 2010) in the R package GAPIT

version 3.0 (Lipka et al., 2012) according to Lin et al. (2020). In brief, the mixed linear model

controlled for maturity, population stratification, and unequal relatedness by including DTA

BLUPs, principal components (PCs) based on a genotype matrix, and a kinship matrix. The

PCs were calculated from the 310 line × 9,715,072 SNP genotype matrix with the 'prcomp'

function in R version 3.5.1 (R core team, 2018). The same SNP genotype matrix was pruned

at a linkage disequilibrium (LD) threshold of $r^2 \leq 0.2$ in PLINK version 1.09_beta5 (Purcell

et al., 2007), resulting in a subset of 287,023 SNPs used to construct the kinship matrix based

on the centered IBS method (Endelman and Jannink, 2012) in TASSEL 5.0 (Bradbury et al.,

2007). The optimal model for GWAS selected by the Bayesian information criterion

(Schwarz, 1978) included only the kinship matrix. The amount of phenotypic variation

explained by a SNP was approximated with the likelihood-ratio-based R^2 statistic (R^2_{LR}) of

Sun et al. (2010).

TWAS

With BLUP values of g_c from the 310 lines as the response variable, TWAS was

performed with PEER values for each of the 20,013 expressed genes by fitting a mixed linear

model using the *gwas* function with the P3D function set to FALSE in the R package

'rrBLUP' version 4.6 (Endelman, 2011). The optimal model included the same kinship matrix from GWAS.

FCT

The FCT was used to combine GWAS and TWAS results according to Kremling et al. (2019). Briefly, the top 10% of the most associated SNPs (971,508) from GWAS were assigned to the nearest gene. The *P*-values of genes not tested in TWAS were set to 1. For each gene, the paired GWAS and TWAS *P*-values were used to conduct a FCT with the sumlog method in the R package 'metap' version 1.4 (Dewey, 2016).

Candidate gene identification

Given that GWAS, TWAS, and FCT have different statistical power and structure, the rankings of *P*-values for each statistical method were used to identify potential candidate genes following that of Kremling et al. (2019). To identify candidate genes from GWAS results, a set of loci was declared among the top 0.01% of SNPs associated with g_c as described by Wu et al. (2021). The search interval for candidate genes was restricted to ± 200 kb of the peak SNP for each of the declared loci following Lin et al. (2020). The top 1% of genes according to their *P*-values were selected from TWAS and FCT results, resulting in comparable numbers of candidate genes across all three methods.

Haplotype-based association analysis of *ISTL1*

Haploblocks in a ± 200 kb region encompassing the candidate gene encoding ISTL1 were constructed using the confidence interval method (Gabriel et al., 2002) in Haploview version 4.2 (Barrett et al., 2005). The identified 91 haploblocks (Supplemental Data Set 3), which were created from an LD pruned ($r^2 > 0.9999$) subset of 726 SNPs, were each tested for an association with g_c by fitting a mixed linear model that included the same GWAS kinship matrix in ASReml-R version 3.0 (Gilmour et al., 2009). Haplotype effects were

28

estimated for each haploblock, and pairwise comparisons were performed using the

predictPlus.asreml function.

Evaluation of NILs for *ISTL1*

Experimental field design

To assess haplotype effects of *ISTL1* on g_c variation in a common genetic background, 13 nested near isogenic lines (nNILs) that contain introgressions harboring *ISTL1* were selected. The evaluated nNILs were derived from crosses between eight diverse inbred lines (CML69, Ki11, Ki3, NC350, NC358, Oh43, Tx303 and Tzi8) and the recurrent inbred parent B73 (Morales et al., 2020). In 2020 at the University of California San Diego, the entire experiment of 13 nNILs and their eight introgression donor parents was planted as a single replicate in a 4 × 6 incomplete block design on two different dates (June 18 and 28) in the same field. Each incomplete block was augmented by B73. Experimental units for the nNIL20 experiment were the same as those in the WiDiv18 experiment.

Analysis of nNILs and their ISTL1 haplotypes

The collection and processing of g_c data from the nNIL20 experiment were similar to the procedures used for the WiDiv18 experiment, with the exception that best linear unbiased estimators (BLUEs; Supplemental Table S9) were calculated. Least-squares means of nNILs and donor lines were compared to B73 with the method of Dunnett (1955) in SAS version 9.4 (SAS institute 2013).

To validate the observed strong effects of *ISTL1* haploblocks on g_c in the WiDiv panel, the haploblocks most significantly associated with g_c in the WiDiv panel were investigated among 12 of the 13 nNILs that produced plants. With HapMap 3.2.1 SNP genotypes in B73 RefGen_v4 coordinates (Bukowski et al., 2018) extracted from *ISTL1* haploblocks, haplotypes for B73 and the donor parents were constructed according to those in

29

the WiDiv panel and projected onto their corresponding nNILs by using information from GBS SNP markers of Morales et al. (2020). The haplotype-based association analysis for g_c was performed as described above, except that the GWAS kinship matrix was not needed to control for genetic background effects among tested nNILs.

Data availability

All raw 3′ mRNA-seq data are available from the NCBI Sequence Read Archive under BioProject PRJNA773975. Supplemental Data Sets 1-3 are available at CyVerse (https://datacommons.cyverse.org/browse/iplant/home/shared/GoreLab/dataFromPubs/Lin_ LeafCuticleTWAS_2021).

Acknowledgements

References

Aharoni A, Dixit S, Jetter R, Thoenes E, van Arkel G, Pereira A (2004) The *SHINE* clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in *Arabidopsis*. Plant Cell **16**: 2463–2480

Anders S, Pyl PT, Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **31**: 166–169

Bakan B, Marion D (2017) Assembly of the cutin polyester: From cells to extracellular cell walls. Plants. doi: 10.3390/plants6040057

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics **21**: 263–265

Bernard A, Domergue F, Pascal S, Jetter R, Renne C, Faure J-D, Haslam RP, Napier JA, Lessire R, Joubès J (2012) Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. Plant Cell **24**: 3106–3118

Bessire M, Borel S, Fabre G, Carraça L, Efremova N, Yephremov A, Cao Y, Jetter R, Jacquat A-C, Métraux J-P, et al (2011) A member of the PLEIOTROPIC DRUG RESISTANCE family of ATP binding cassette transporters is required for the formation of a functional cuticle in *Arabidopsis*. Plant Cell **23**: 1958–1970

Bi H, Luang S, Li Y, Bazanova N, Morran S, Song Z, Perera MA, Hrmova M, Borisjuk N, Lopato S (2016) Identification and characterization of wheat drought-responsive MYB transcription factors involved in the regulation of cuticle biosynthesis. J Exp Bot **67**: 5363–5380

Bird D, Beisson F, Brigham A, Shin J, Greer S, Jetter R, Kunst L, Wu X, Yephremov A, Samuels L (2007) Characterization of *Arabidopsis* ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. Plant J **52**: 485–498

Borisjuk N, Hrmova M, Lopato S (2014) Transcriptional regulation of cuticle biosynthesis. Biotechnol Adv **32**: 526–540

Bourdenx B, Bernard A, Domergue F, Pascal S, Léger A, Roby D, Pervent M, Vile D, Haslam RP, Napier JA, et al (2011) Overexpression of *Arabidopsis ECERIFERUM1* promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. Plant Physiol **156**: 29–45

Bourgault R, Matschi S, Vasquez M, Qiao P, Sonntag A, Charlebois C, Mohammadi M, Scanlon MJ, Smith LG, Molina I (2020) Constructing functional cuticles: Analysis of relationships between cuticle lipid composition, ultrastructure and water barrier function in developing adult maize leaves. Annals of Botany. doi: 10.1093/aob/mcz143

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics **23**: 2633–2635

Browning BL, Zhou Y, Browning SR (2018) A One-Penny Imputed Genome from Next-Generation Reference Panels. Am J Hum Genet **103**: 338–348

Bukowski R, Guo X, Lu Y, Zou C, He B, Rong Z, Wang B, Xu D, Yang B, Xie C, et al (2018) Construction of the third-generation *Zea mays* haplotype map. GigaScience **7**: 1–12

Buono RA, Paez-Valencia J, Miller ND, Goodman K, Spitzer C, Spalding EP, Otegui MS (2016) Role of SKD1 regulators LIP5 and IST1-LIKE1 in endosomal sorting and plant development. Plant Physiol **171**: 251–264

Busta L, Schmitz E, Kosma DK, Schnable JC, Cahoon EB (2021) A co-opted steroid synthesis gene, maintained in sorghum but not maize, is associated with a divergence in leaf wax chemistry. Proc Natl Acad Sci U S A. doi: 10.1073/pnas.2022982118

Chen G, Komatsuda T, Ma JF, Nawrath C, Pourkheirandish M, Tagiri A, Hu Y-G, Sameri M, Li X, Zhao X, et al (2011) An ATP-binding cassette subfamily G full transporter is essential for the retention of leaf water in both wild barley and rice. Proc Natl Acad Sci USA **108**: 12354–12359

Chen YA, Scheller RH (2001) SNARE-mediated membrane fusion. Nat Rev Mol Cell Biol **2**: 98–106

DeBono A, Yeats TH, Rose JKC, Bird D, Jetter R, Kunst L, Samuels L (2009) *Arabidopsis* LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. Plant Cell **21**: 1230–1238

Dewey M (2016) metap: Meta-analysis of significance values. R package version 0.7.

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in *Arabidopsis*. Trends Plant Sci **15**: 573–581

Dunnett CW (1955) A Multiple Comparison Procedure for Comparing Several Treatments with a Control. J Am Stat Assoc **50**: 1096–1121

Elango D, Xue W, Chopra S (2020) Genome wide association mapping of epi-cuticular wax genes in Sorghum bicolor. Physiol Mol Biol Plants **26**: 1727–1737

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome **4**: 250–255

Endelman JB, Jannink J-L (2012) Shrinkage estimation of the realized relationship matrix. G3 **2**: 1405–1413

Febrero A, Fernandez S, Molina-Cano JL, Araus JL (1998) Yield, carbon isotope discrimination, canopy reflectance and cuticular conductance of barley isolines of

differing glaucousness. J Exp Bot **49**: 1575–1581

Ferguson JN, Fernandes SB, Monier B, Miller ND, Allen D, Dmitrieva A, Schmuker P, Lozano R, Valluru R, Buckler ES, et al (2021) Machine learning-enabled phenotyping for GWAS and TWAS of WUE traits in 869 field-grown sorghum accessions. Plant Physiol. doi: 10.1093/plphys/kiab346

Fich EA, Segerson NA, Rose JKC (2016) The plant polyester cutin: biosynthesis, structure, and biological roles. Annu Rev Plant Biol **67**: 207–233

Franke R, Höfer R, Briesen I, Emsermann M, Efremova N, Yephremov A, Schreiber L (2009) The *DAISY* gene from *Arabidopsis* encodes a fatty acid elongase condensing enzyme involved in the biosynthesis of aliphatic suberin in roots and the chalaza-micropyle region of seeds. Plant J **57**: 80–95

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, et al (2002) The structure of haplotype blocks in the human genome. Science **296**: 2225–2229

Gil HL, Kim J, Chung MS, Joon PS (2017) The MIEL1 E3 Ubiquitin Ligase Negatively Regulates Cuticular Wax Biosynthesis in *Arabidopsis* Stems. Plant Cell Physiol **58**: 1249–1259

Gilmour AR, Gogel BJ, Cullis BR, Thompson R, Butler D, Others (2009) ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK

Girard A-L, Mounet F, Lemaire-Chamley M, Gaillard C, Elmorjani K, Vivancos J, Runavot J-L, Quemener B, Petit J, Germain V, et al (2012) Tomato GDSL1 is required for cutin deposition in the fruit cuticle. Plant Cell **24**: 3119–3134

Goodman K, Paez-Valencia J, Pennington J, Sonntag A, Ding X, Lee HN, Ahlquist PG, Molina I, Otegui MS (2021) ESCRT components ISTL1 and LIP5 are required for tapetal function and pollen viability. The Plant Cell. doi: 10.1093/plcell/koab132

Guo J, Xu W, Yu X, Shen H, Li H, Cheng D, Liu A, Liu J, Liu C, Zhao S, et al (2016) Cuticular wax accumulation is associated with drought tolerance in wheat near-isogenic lines. Front Plant Sci **7**: 1809

Guo W, Roth D, Walch-Solimena C, Novick P (1999) The exocyst is an effector for Sec4p, targeting secretory vesicles to sites of exocytosis. EMBO J **18**: 1071–1080

Hansey CN, Johnson JM, Sekhon RS, Kaeppler SM, de Leon N (2011) Genetic diversity of a maize association population with restricted phenology. Crop Science **51**: 704

Hen-Avivi S, Savin O, Racovita RC, Lee W-S, Adamski NM, Malitsky S, Almekias-Siegl E, Levy M, Vautrin S, Bergès H, et al (2016) A metabolic gene cluster in the wheat *W1* and the barley *Cer-cqu* loci determines β-diketone biosynthesis and glaucousness. Plant Cell **28**: 1440–1460

Hill CP, Babst M (2012) Structure and function of the membrane deformation AAA ATPase

Vps4. Biochim Biophys Acta **1823**: 172–181

Hirsch CN, Foerster JM, Johnson JM, Sekhon RS, Muttoni G, Vaillancourt B, Peñagaricano F, Lindquist E, Pedraza MA, Barry K, et al (2014) Insights into the maize pan-genome and pan-transcriptome. Plant Cell **26**: 121–135

Hooker TS, Lam P, Zheng H, Kunst L (2007) A core subunit of the RNA-processing/degrading exosome specifically influences cuticular wax biosynthesis in *Arabidopsis*. Plant Cell **19**: 904–913

Hunt L, Amsbury S, Baillie A, Movahedi M, Mitchell A, Afsharinafar M, Swarup K, Denyer T, Hobbs JK, Swarup R, et al (2017) Formation of the Stomatal Outer Cuticular Ledge Requires a Guard Cell Wall Proline-Rich Protein. Plant Physiol **174**: 689–699

Jiang Y, Duan Y, Yin J, Ye S, Zhu J, Zhang F, Lu W, Fan D, Luo K (2014) Genome-wide identification and characterization of the Populus WRKY transcription factor family and analysis of their expression in response to biotic and abiotic stresses. J Exp Bot **65**: 6629–6644

Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin C-S, et al (2017) Improved maize reference genome with single-molecule technologies. Nature **546**: 524–527

Jouannic N, Lepetit M, Vergnolle C, Cantrel C, Gardies AM, Kader JC, Arondel V (1998) Isolation of a cDNA from *Arabidopsis thaliana* that complements the *sec14* mutant of yeast. Eur J Biochem **258**: 402–410

Kamiya T, Borghi M, Wang P, Danku JMC, Kalmbach L, Hosmani PS, Naseer S, Fujiwara T, Geldner N, Salt DE (2015) The MYB36 transcription factor orchestrates Casparian strip formation. Proc Natl Acad Sci U S A **112**: 10533–10538

Kant P, Gordon M, Kant S, Zolla G, Davydov O, Heimer YM, Chalifa-Caspi V, Shaked R, Barak S (2008) Functional-genomics-based identification of genes that regulate *Arabidopsis* responses to multiple abiotic stresses. Plant Cell Environ **31**: 697–714

Kerstiens G (2006) Water transport in plant cuticles: an update. J Exp Bot **57**: 2493–2499

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol **37**: 907–915

Kim H, Lee SB, Kim HJ, Min MK, Hwang I, Suh MC (2012) Characterization of glycosylphosphatidylinositol-anchored lipid transfer protein 2 (LTPG2) and overlapping function between LTPG/LTPG1 and LTPG2 in cuticular wax export or accumulation in *Arabidopsis thaliana*. Plant Cell Physiol **53**: 1391–1403

Kosma DK, Bourdenx B, Bernard A, Parsons EP, Lü S, Joubès J, Jenks MA (2009) The impact of water deficiency on leaf cuticle lipids of *Arabidopsis*. Plant Physiol **151**: 1918–1929

Kosma DK, Jenks MA (2007) Eco-physiological and molecular-genetic determinants of

plant cuticle function in drought and salt stress tolerance. Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops 91–120

Kremling KAG, Diepenbrock CH, Gore MA, Buckler ES, Bandillo NB (2019) Transcriptome-Wide Association Supplements Genome-Wide Association in *Zea mays*. G3 **9**: 3023–3033

Lee C, Goldberg J (2010) Structure of coatomer cage proteins and the relationship among COPI, COPII, and clathrin vesicle coats. Cell **142**: 123–132

Lee HG, Seo PJ (2016) The *Arabidopsis* MIEL1 E3 ligase negatively regulates ABA signalling by promoting protein turnover of MYB96. Nat Commun **7**: 12525

Lee S-B, Jung S-J, Go Y-S, Kim H-U, Kim J-K, Cho H-J, Park OK, Suh M-C (2009) Two *Arabidopsis* 3-ketoacyl CoA synthase genes, *KCS20* and *KCS2/DAISY,* are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. The Plant Journal **60**: 462–475

Lee SB, Suh MC (2015) Advances in the understanding of cuticular waxes in *Arabidopsis thaliana* and crop species. Plant Cell Rep **34**: 557–572

Li L, Du Y, He C, Dietrich CR, Li J, Ma X, Wang R, Liu Q, Liu S, Wang G, et al (2019) Maize *glossy6* is involved in cuticular wax deposition and drought tolerance. J Exp Bot **70**: 3089–3099

Lin M, Matschi S, Vasquez M, Chamness J, Kaczmar N, Baseggio M, Miller M, Stewart EL, Qiao P, Scanlon MJ, et al (2020) Genome-Wide Association Study for Maize Leaf Cuticular Conductance Identifies Candidate Genes Involved in the Regulation of Cuticle Development. G3 . doi: 10.1534/g3.119.400884

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics **28**: 2397–2399

Lodish H, Berk A, Kaiser CA, Krieger M, Bretscher A, Ploegh H, Amon A, Martin KC (2016) Molecular Biology of the Cell, 8th edition. W.H. Freeman, New York.

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**: 550

Lü S, Zhao H, Des Marais DL, Parsons EP, Wen X, Xu X, Bangarusamy DK, Wang G, Rowland O, Juenger T, et al (2012) *Arabidopsis ECERIFERUM9* involvement in cuticle formation and maintenance of plant water status. Plant Physiol **159**: 930–944

Marino D, Froidure S, Canonne J, Ben Khaled S, Khafif M, Pouzet C, Jauneau A, Roby D, Rivas S (2013) *Arabidopsis* ubiquitin ligase MIEL1 mediates degradation of the transcription factor MYB30 weakening plant defence. Nature Communications. doi: 10.1038/ncomms2479

Matschi S, Vasquez MF, Bourgault R, Steinbach P, Chamness J, Kaczmar N, Gore MA,

Molina I, Smith LG (2020) Structure‑function analysis of the maize bulliform cell cuticle and its potential role in dehydration and leaf rolling. Plant Direct. doi: 10.1002/pld3.282

Mayers JR, Hu T, Wang C, Cárdenas JJ, Tan Y, Pan J, Bednarek SY (2017) SCD1 and SCD2 Form a Complex That Functions with the Exocyst and RabE1 in Exocytosis and Cytokinesis. Plant Cell **29**: 2610–2625

McFarlane HE, Shin JJH, Bird DA, Samuels AL (2010) *Arabidopsis* ABCG transporters, which are required for export of diverse cuticular lipids, dimerize in different combinations. Plant Cell **22**: 3066–3075

McFarlane HE, Watanabe Y, Yang W, Huang Y, Ohlrogge J, Lacey Samuels A (2014) Golgi- and trans-golgi network-mediated vesicle trafficking is required for wax secretion from epidermal cells. Plant Physiol **164**: 1250–1260

Morales L, Repka AC, Swarts KL, Stafstrom WC, He Y, Sermons SM, Yang Q, Lopez‑Zuniga LO, Rucker E, Thomason WE, et al (2020) Genotypic and phenotypic characterization of a large, diverse population of maize near‑isogenic lines. Plant J. doi: 10.1111/tpj.14787

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W (1996) Applied linear statistical models. Irwin Chicago

Oshima Y, Mitsuda N (2013) The MIXTA-like transcription factor MYB16 is a major regulator of cuticle formation in vegetative organs. Plant Signal Behav **8**: e26826

Panikashvili D, Shi JX, Bocobza S, Franke RB, Schreiber L, Aharoni A (2010) The *Arabidopsis* DSO/ABCG11 transporter affects cutin metabolism in reproductive organs and suberin in roots. Mol Plant **3**: 563–575

Patwari P, Salewski V, Gutbrod K, Kreszies T, Dresen-Scholz B, Peisker H, Steiner U, Meyer AJ, Schreiber L, Dörmann P (2019) Surface wax esters contribute to drought tolerance in *Arabidopsis*. Plant J **98**: 727–744

Philippe G, Geneix N, Petit J, Guillon F, Sandt C, Rothan C, Lahaye M, Marion D, Bakan B (2020) Assembly of tomato fruit cuticles: a cross-talk between the cutin polyester and cell wall polysaccharides. New Phytol **226**: 809–822

Pighin JA, Zheng H, Balakshin LJ, Goodman IP, Western TL, Jetter R, Kunst L, Samuels AL (2004) Plant cuticular lipid export requires an ABC transporter. Science **306**: 702–704

Pignon CP, Fernandes SB, Valluru R, Bandillo N, Lozano R, Buckler E, Gore MA, Long SP, Brown PJ, Leakey ADB (2021) Phenotyping stomatal closure by thermal imaging for GWAS and TWAS of water use efficiency-related genes. Plant Physiol. doi: 10.1093/plphys/kiab395

Pollard M, Beisson F, Li Y, Ohlrogge JB (2008) Building lipid barriers: biosynthesis of cutin and suberin. Trends Plant Sci **13**: 236–246

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet **81**: 559–575

Qiao P, Bourgault R, Mohammadi M, Matschi S, Philippe G, Smith LG, Gore MA, Molina I, Scanlon MJ (2020) Transcriptomic network analyses shed light on the regulation of cuticle development in maize leaves. Proc Natl Acad Sci U S A **117**: 12464–12471

Qiao P, Bourgault R, Mohammadi M, Smith LG, Gore MA, Molina I, Scanlon MJ (2019) Network analyses implicate a role for PHYTOCHROME-mediated light signaling in the regulation of cuticle development in plant leaves. bioRxiv 812107

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Raffaele S, Vailleau F, Léger A, Joubès J, Miersch O, Huard C, Blée E, Mongrand S, Domergue F, Roby D (2008) A MYB transcription factor regulates very-long-chain fatty acid biosynthesis for activation of the hypersensitive cell death response in *Arabidopsis*. Plant Cell **20**: 752–767

Riederer M, Schreiber L (2001) Protecting against water loss: analysis of the barrier properties of plant cuticles. J Exp Bot **52**: 2023–2032

Rincent R, Laloë D, Nicolas S, Altmann T, Brunel D, Revilla P, Rodríguez VM, Moreno-Gonzalez J, Melchinger A, Bauer E, et al (2012) Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays L.*). Genetics **192**: 715–728

Saito C, Ueda T (2009) Chapter 4: functions of RAB and SNARE proteins in plant life. Int Rev Cell Mol Biol **274**: 183–233

Schwarz G (1978) Estimating the Dimension of a Model. Ann Stat **6**: 461–464

Seo PJ, Lee SB, Suh MC, Park M-J, Go YS, Park C-M (2011) The MYB96 transcription factor regulates cuticular wax biosynthesis under drought conditions in *Arabidopsis*. Plant Cell **23**: 1138–1152

Shepherd T, Wynne Griffiths D (2006) The effects of stress on plant cuticular waxes. New Phytol **171**: 469–499

Stegle O, Parts L, Durbin R, Winn J (2010) A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. PLoS Comput Biol **6**: e1000770

Stenmark H (2009) Rab GTPases as coordinators of vesicle traffic. Nat Rev Mol Cell Biol **10**: 513–525

Strobl C, Boulesteix A-L, Zeileis A, Hothorn T (2007) Bias in random forest variable

importance measures: illustrations, sources and a solution. BMC Bioinformatics **8**: 25

Sturaro M, Hartings H, Schmelzer E, Velasco R, Salamini F, Motto M (2005) Cloning and characterization of *GLOSSY1*, a maize gene involved in cuticle membrane and wax production. Plant Physiol **138**: 478–489

Sun G, Zhu C, Kramer MH, Yang S-S, Song W, Piepho H-P, Yu J (2010) Variation explained in mixed-model association mapping. Heredity **105**: 333–340

Tibbs Cortes L, Zhang Z, Yu J (2021) Status and prospects of genome-wide association studies in plants. Plant Genome **14**: e20077

Wang W, Zhang Y, Xu C, Ren J, Liu X, Black K, Gai X, Wang Q, Ren H (2015) Cucumber *ECERIFERUM1* (*CsCER1*), which influences the cuticle properties and drought tolerance of cucumber, plays a key role in VLC alkanes biosynthesis. Plant Molecular Biology **87**: 219–233

Wang Y, Jin S, Xu Y, Li S, Zhang S, Yuan Z, Li J, Ni Y (2020) Overexpression of *BnKCS1-1*, *BnKCS1-2*, and *BnCER1-2* promotes cuticular wax production and increases drought tolerance in *Brassica napus*. The Crop Journal **8**: 26–37

Wang Y, Wan L, Zhang L, Zhang Z, Zhang H, Quan R, Zhou S, Huang R (2012) An ethylene response factor OsWR1 responsive to drought stress transcriptionally activates wax synthesis related genes and increases wax production in rice. Plant Mol Biol **78**: 275–288

Wilmer C, Fricker M (1996) Stomata, 2nd edition.

Wu D, Tanaka R, Li X, Ramstein GP, Cu S, Hamilton JP, Buell CR, Stangoulis J, Rocheford T, Gore MA (2021) High-resolution genome-wide association study pinpoints metal transporter and chelator genes involved in the genetic control of element levels in maize grain. G3 . doi: 10.1093/g3journal/jkab059

Xue D, Zhang X, Lu X, Chen G, Chen Z-H (2017) Molecular and evolutionary mechanisms of cuticular wax for plant drought tolerance. Front Plant Sci **8**: 621

Yeats TH, Martin LBB, Viart HM-F, Isaacson T, He Y, Zhao L, Matas AJ, Buda GJ, Domozych DS, Clausen MH, et al (2012) The identification of cutin synthase: formation of the plant polyester cutin. Nat Chem Biol **8**: 609–611

Yeats TH, Rose JKC (2013) The formation and function of plant cuticles. Plant Physiol **163**: 5–20

Zhang J-Y, Broeckling CD, Blancaflor EB, Sledge MK, Sumner LW, Wang Z-Y (2005) Overexpression of *WXP1*, a putative *Medicago truncatula* AP2 domain-containing transcription factor gene, increases cuticular wax accumulation and enhances drought tolerance in transgenic alfalfa (*Medicago sativa*). Plant J **42**: 689–707

Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, et al (2010) Mixed linear model approach adapted for

genome-wide association studies. Nat Genet **42**: 355–360

Zhao L, Sack FD (1999) Ultrastructure of stomatal development in *Arabidopsis* (Brassicaceae) leaves. Am J Bot **86**: 929–939

Zhou L, Ni E, Yang J, Zhou H, Liang H, Li J, Jiang D, Wang Z, Liu Z, Zhuang C (2013) Rice *OsGL1-6* is involved in leaf cuticular wax accumulation and drought resistance. PLoS One **8**: e65139

Zhu X, Xiong L (2013) Putative megaenzyme DWA1 plays essential roles in drought resistance by regulating stress-induced wax deposition in rice. Proc Natl Acad Sci USA **110**: 17790–17795

Figure 1.

Box-plot of importance scores for cuticular wax composition and structural features in random forest regression to predict maize leaf cuticular conductance (g_c) using five-fold cross validation for 50 times.

Figure 2.

Manhattan plot of results from a genome-wide association study (A; GWAS), a transcriptome-wide association study (B; TWAS), and Fisher's combined test (C; FCT) of adult maize leaf cuticular conductance (g_c) across two environments in San Diego, CA. (A) The -$\log_{10}$ P-value of each SNP tested in a mixed linear model analysis of g_c is plotted as a point against its physical position (B73 RefGen_v4) for the 10 chromosomes of maize. The peak SNPs closest to plausible candidate genes are colored based on their functional categories. (B) The -$\log_{10}$ P-value of each gene transcript tested in a mixed linear model analysis of g_c is plotted as a point against its physical position (B73 RefGen_v4). Plausible candidate genes are colored based on their functional categories. (C) The -$\log_{10}$ P-value of each of the top 10% SNPs in GWAS paired with the nearest gene in FCT plotted as a point against its physical position (B73 RefGen_v4). The most significant SNP-gene pairs related to plausible candidate genes identified in the FCT are colored based on their functional categories.

Figure 3.

Manhattan plot of results from a haplotype-based association analysis in the vicinity of *ISTL1* for adult maize leaf cuticular conductance (g_c) combined from two environments in San Diego, CA, in the Wisconsin diversity panel. The -$\log_{10}$ P-value of each haploblock tested in a mixed linear model analysis of g_c is plotted as a short horizontal line against its physical position (B73 RefGen_v4). The length of a short horizontal line corresponds to the size of

that haplotype block. The grey vertical bar represents the physical position of *ISTL1*. The top two haploblocks (Blocks 65 and 52) that are most associated with g_c are indicated. Polymorphic and monomorphic haploblocks among the evaluated nested near-isogenic lines are colored in blue and red, respectively.

Figure 4.

Scatter plot of best linear unbiased estimator values for adult maize leaf cuticular conductance (g_c) for B73, nested near-isogenic lines (nNILs) containing *ISTL1* introgressions and their donor parent lines across two environments in San Diego, CA, in 2020. Lines with g_c significantly different from B73 are highlighted in orange. '*' represents *P*-value < 0.1 and '**' represents *P*-value < 0.05 after adjustment by the Dunnett's method.

Table 1.

Plausible candidate genes identified in a genome-wide association study (GWAS), a transcriptome-wide association study (TWAS) and the Fisher's combined test (FCT) for leaf cuticular conductance (g_c) in the Wisconsin diversity panel.

Supplemental information

Supplemental Figure S1.

Guard cell cuticle structural features analyzed for relationship to cuticular conductance. Transverse sections of formaldehyde-acetic acid-alcohol-fixed, cryosectioned leaves were stained with Fluorol Yellow to fluorescently label the cuticle and imaged via superresolution confocal microscopy as described by Matschi et al. 2020. Images were collected at a position approximately halfway from base to tip of a guard cell pair (GC, guard cells; F, cuticle flap; B, bumps lining the inner surfaces of the guard cell pore). Note the semi-complementary pattern of bumps on adjacent GC surfaces. Structural features analyzed for correlation with cuticular conductance were defined as illustrated schematically. Roundness is the ratio between the length of a straight line measuring GC height (yellow line) and a curved line tracing the inner surface of the guard cell pore without regard to bumps (red). Bumpiness is the ratio between the length of a line tracing the bumps on the GC pore surface (green line) and a curved line tracing the guard cell pore surface without regard to bumps (red). Flap length is the length of a cuticular flap (blue line). Sealability attempts to measure the capacity of adjacent GC flaps to overlap when the GC pore is closed in the presence of variability in our image collection in the degree of stomatal closure. It was measured as the difference in length between a line connecting the tips of GC flaps (orange line 1) and a line connecting the inner surfaces of adjacent guard cell pores (orange line 2).

Supplemental Figure S2.

Upset plot showing the number of overlapping plausible candidate genes between a genome-wide association study (GWAS), a transcriptome-wide association study (TWAS), and the Fisher's combined test (FCT).

42

Supplemental Figure S3.

Haplotype effect analysis of Block 52 in the Wisconsin diversity (WiDiv) panel and nested near-isogenic lines (nNILs). (A) Haplotype effects of Block 52 estimated in a mixed linear model analysis of adult maize leaf cuticular conductance (g_c) in the WiDiv panel. The three haplotypes (1, 5, and 7) segregating among the nNILs are in bolded font. (B) The effects of the three segregating haplotypes estimated with the nNILs. The upper triangle of the matrix shows differences between haplotype effects, whereas the lower triangle shows the significance of the differences as *P*-values. (C) Pearson's correlation between haplotype effects estimated in the WiDiv panel and nNILs.

Supplemental Table S1.

Relationship of g_c to cuticular wax composition and structural features. Included in the table are importance scores from random forest regression and Pearson's correlation coefficients between best linear unbiased predictors for g_c and averaged observations for cuticular wax composition and structural features.

Supplemental Table S2.

Genomic information (B73 RefGen_v4) for the top 1% genes associated with maize leaf cuticular conductance (g_c) in Fisher's combined test in the Wisconsin Diversity panel.

Supplemental Table S3.

Genomic information (B73 RefGen_v4) and association statistics for the top 0.01% SNPs associated with maize leaf cuticular conductance (g_c) in a genome-wide association study in the Wisconsin Diversity panel.

Supplemental Table S4.

Genomic information (B73 RefGen_v4) for the candidate genes residing within ± 200 kb of the peak SNPs identified in the genome-wide association study in the Wisconsin diversity panel.

Supplemental Table S5.

Genomic information (B73 RefGen_v4) for the top 1% associated genes for leaf cuticular conductance (g_c) in the transcriptome-wide association study in the Wisconsin diversity panel.

Supplemental Table S6.

Rice and Arabidopsis homologs of 22 plausible candidate genes identified for leaf cuticular conductance (g_c) via a genome-wide association study, transcriptome-wide association analysis and Fisher's combined test in the maize Wisconsin diversity panel.

Supplemental Table S7.

Statistics of associations between g_c and 91 haploblocks in the vicinity of *ISTL1* in the Wisconsin diversity panel.

Supplemental Table S8.

Genomic introgressions for nested near-isogenic lines (nNILs) containing the *ISTL1* genomic region. Included in the table are physical positions and flanking markers for these introgressions.

Supplemental Table S9.

Best linear unbiased estimator (BLUE) values of adult maize leaf cuticular conductance (g_c; $g·h^{-1}·g^{-1}$) and differences of least squares means (LSM) of g_c for nested near-isogenic lines (nNILs) containing *ISTL1* introgressions and their donor parents from two environments in San Diego, CA, in 2020.

Supplemental Table S10.

Best fitted models used to calculate best linear unbiased predictors for adult maize leaf cuticular conductance (g_c) and flowering time (days to anthesis) for two environments for the Wisconsin Diversity (WiDiv) panel, as well as best linear unbiased estimators for g_c for two environments for the nested near-isogenic lines (nNILs), according to a likelihood ratio test ($\alpha = 0.05$). The asterisk (*) indicates that a random effect term was retained in the mixed linear model, whereas the 'x' indicates that a fitted random effect term was not significant and removed from the mixed linear model.

Supplemental Table S11.

Best linear unbiased predictors (BLUPs) of adult maize leaf cuticular conductance (g_c; $g·h^{-1}·g^{-1}$), BLUPs of flowering time (FT; days to anthesis) for 310 maize inbred lines from combined environments in San Diego, CA, in 2018.

Supplemental Table S12.

Cuticular waxes and stomatal cuticle features potentially associated with cuticular conductance (g_c) for lines consistently at the upper or lower end of the range for the g_c phenotype in San Diego (2016 and 2017; Lin et al. 2020). Included in the table are group classification of lines following assignments of Hansey et al. (2011), best linear unbiased predictors (BLUPs) of adult maize leaf g_c ($g \cdot h^{-1} \cdot g^{-1}$), extreme phenotype classification (high or low g_c), and averaged abundance of cuticular waxes (free alcohols, free fatty acids, hydrocarbons, aldehydes, wax esters and alicyclics) and stomatal cuticle features (sealability, flap length and bumpiness) from two environments in San Diego in 2018.